

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:25:06 ; Search time 50 Seconds
(without alignments)
1812.657 Million cell updates/sec

Title: US-09-715-927-6

Perfect score: 2929

Sequence: 1 MTRAGDHNRQRCGLADY.....ACGPAKEVRKENQANTSVW 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_13Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2929	100.0	571	22 ABUS2922	Cell structure and
2	2929	100.0	571	22 AAM39134	Human polypeptide
3	2929	100.0	571	22 AAB27245	Human EXMAD-23 SEQ
4	2925	99.9	571	23 AAE23866	Human ferroportin
5	2917	99.6	571	20 AAY29327	Human secreted pro
6	2917	99.6	571	22 AAU39053	Human secreted pro
7	2917	99.6	571	23 ABB55762	Human polypeptide
8	2018	68.9	408	23 ABP41875	Human ovarian anti
9	1789	61.1	382	23 ABB89115	Human polypeptide

10	1789	61.1	383	19 AAW74869	Human secreted pro
11	1789	61.1	383	23 ABG95320	Human novel secret
12	1630	55.7	439	22 AAM40920	Human polypeptide
13	1611	55.0	439	22 ABG20262	Novel human diagno
14	1470.5	50.2	319	22 ABUS2923	Cell structure and
15	859	29.3	167	22 AAB93563	Human protein sequ
16	606	20.7	123	21 AAG00622	Human secreted pro
17	302	10.3	576	23 ABB93606	Herbicidally activ
18	295	10.1	150	24 ABG75821	Transporters and i
19	181.5	6.2	371	21 AAG09603	Arabidopsis thalia
20	159	5.4	242	21 AAG09604	Arabidopsis thalia
21	157	5.4	496	23 ABP65787	Bifidobacterium lo
22	143	4.9	231	21 AAG09605	Arabidopsis thalia
23	128	4.4	1313	22 ABB63904	Drosophila melanog
24	123.5	4.2	524	23 ABG76953	Human protein, hom
25	123.5	4.2	524	23 ABB05602	Human glucose tran
26	120	4.1	421	23 ABP66017	Bifidobacterium lo
27	120	4.1	458	22 AAU03207	Fruit fly G protei
28	119	4.1	537	22 AAG92976	C glutamicum prote
29	115.5	3.9	491	22 ABB70144	Drosophila melanog
30	115.5	3.9	683	22 AAU44669	Propionibacterium
31	115	3.9	538	22 AAG83020	S. epidermidis ope
32	114.5	3.9	424	22 AAG89887	C glutamicum prote
33	114.5	3.9	424	22 AAB78954	C. glutamicum SR
34	114.5	3.9	424	22 AAB76809	Corynebacterium gl
35	114	3.9	392	21 AAY81756	Streptococcus pneu
36	114	3.9	392	24 ABU01535	S. pneumoniae type
37	113.5	3.9	373	19 AAM80675	S. pneumoniae prot
38	113.5	3.9	522	18 AAM7836	Rat glucose transp
39	113	3.9	475	22 ABB57889	Drosophila melanog
40	113	3.9	522	22 AAB86970	D. melanogaster pe
41	112	3.8	475	22 ABB67155	Drosophila melanog
42	111.5	3.8	512	23 AAO14190	Human transporter
43	109.5	3.7	470	23 ABP39765	Staphylococcus epi
44	109.5	3.7	480	21 AAG28044	Arabidopsis thalia
45	109.5	3.7	483	21 AAG28043	Arabidopsis thalia

ALIGNMENTS

RESULT 1

ABUS2922
ID ABUS2922 standard; Protein; 571 AA.

XX AC ABUS2922;

XX DT 15-APR-2003 (first entry)

XX DE Cell structure and motility-associated protein from DKFZphutel_24j6.

XX KW Human; gene therapy; vaccine; disease treatment; detection.

XX OS Homo sapiens.

XX PN WO200112659-A2.

XX PD 22-FEB-2001.

XX PF 18-AUG-2000; 2000WO-IB01496.

XX PR 18-AUG-1999; 99US-0149499.

XX PR 28-SEP-1999; 99US-0156503.

XX (GBHU-) GERMAN HUMAN GENOME PROJECT.

XX WIemann S;

XX WPI; 2001-327840/34.

XX N-PSDB; ABX71330.

XX Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies -

XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

Claim 21; Page 537; 1095pp; English.

This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence represents a polypeptide described in the disclosure of the invention.

Query Match 100.0%; Score 2929; DB 22; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.7e-271;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTRAGDHNQRQCGCGSLADYLTSAKFLLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60
DB 1 MTRAGDHNQRQCGCGSLADYLTSAKFLLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60

QY 61 TAVYGLVAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKHEL 120
DB 61 TAVYGLVAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKHEL 120

QY 121 LTMVHGWLTSYLLITIANIANLASTATAITIQRDWIVVVGEDRSKLANNNATIRRI 180
DB 121 LTMVHGWLTSYLLITIANIANLASTATAITIQRDWIVVVGEDRSKLANNNATIRRI 180

QY 181 DQLTNIIAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYLLMKVYQKTPALAVKAGLK 240
DB 181 DQLTNIIAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYLLMKVYQKTPALAVKAGLK 240

QY 241 EEETELQNLNKHDEPKPLEGTHLMGVKDSNIHELEHEOEPTCASOMAPPERTFDGW 300
DB 241 EEETELQNLNKHDEPKPLEGTHLMGVKDSNIHELEHEOEPTCASOMAPPERTFDGW 300

QY 301 SYNQPVFLAGMLAFLYMTVLGFDCTTGAYYTGQISGLSILMGASAITGIMGTVA 360
DB 301 SYNQPVFLAGMLAFLYMTVLGFDCTTGAYYTGQISGLSILMGASAITGIMGTVA 360

QY 361 TWLRKGLVRTGLISGLAQSLILCVISVFMGSPDLDSVPEDIRSRFTQGESITP 420
DB 361 TWLRKGLVRTGLISGLAQSLILCVISVFMGSPDLDSVPEDIRSRFTQGESITP 420

QY 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLWSFDLTVTQL 480
DB 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLWSFDLTVTQL 480

QY 481 QENVIESERGIINGVQSNMYLLDLHFIIMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
DB 481 QENVIESERGIINGVQSNMYLLDLHFIIMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540

QY 541 FAQNTLGNLKFACGPDPAKEVRKENQANTSV 571
DB 541 FAQNTLGNLKFACGPDPAKEVRKENQANTSV 571

RESULT 2
AAM39134
ID AAM39134 standard; Protein; 571 AA.
XX
AC
XX
XX
DT
XX
DE
XX
KW
KW
KW

22-OCT-2001 (first entry)
Human polypeptide SEQ ID NO 2279.
Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;

Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.
Homo sapiens.
WO200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US34263.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;
WPI; 2001-442253/47.
N-PSDB; AAI58290.
Novel nucleic acids and polypeptides, useful for treating disorders
such as central nervous system injuries -
Example 4; SEQ ID NO 2279; 10078pp; English.
The invention relates to human nucleic acids (AAI57798-AAI61369) and
the encoded polypeptides (AAM38642-AAM42213) with nontropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
centralised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, anyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders.
Note: The sequence data for this patent did not form part of the printed
specification.

Query Match 100.0%; Score 2929; DB 22; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.7e-271;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTRAGDHNQRQCGCGSLADYLTSAKFLLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60
DB 1 MTRAGDHNQRQCGCGSLADYLTSAKFLLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60

QY 61 TAVYGLVAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKHEL 120
DB 61 TAVYGLVAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKHEL 120

QY 121 LTMVHGWLTSYLLITIANIANLASTATAITIQRDWIVVVGEDRSKLANNNATIRRI 180
DB 121 LTMVHGWLTSYLLITIANIANLASTATAITIQRDWIVVVGEDRSKLANNNATIRRI 180

QY 181 DQLTNIIAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYLLMKVYQKTPALAVKAGLK 240
DB 181 DQLTNIIAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYLLMKVYQKTPALAVKAGLK 240

QY 241 EEETELQNLNKHDEPKPLEGTHLMGVKDSNIHELEHEOEPTCASOMAPPERTFDGW 300
DB 241 EEETELQNLNKHDEPKPLEGTHLMGVKDSNIHELEHEOEPTCASOMAPPERTFDGW 300

QY 301 SYNQPVFLAGMLAFLYMTVLGFDCTTGAYYTGQISGLSILMGASAITGIMGTVA 360
DB 301 SYNQPVFLAGMLAFLYMTVLGFDCTTGAYYTGQISGLSILMGASAITGIMGTVA 360

QY 361 TWLRKGLVRTGLISGLAQSLILCVISVFMGSPDLDSVPEDIRSRFTQGESITP 420
DB 361 TWLRKGLVRTGLISGLAQSLILCVISVFMGSPDLDSVPEDIRSRFTQGESITP 420

QY 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLWSFDLTVTQL 480
DB 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLWSFDLTVTQL 480

QY 481 QENVIESERGIINGVQSNMYLLDLHFIIMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
DB 481 QENVIESERGIINGVQSNMYLLDLHFIIMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540

QY 541 FAQNTLGNLKFACGPDPAKEVRKENQANTSV 571
DB 541 FAQNTLGNLKFACGPDPAKEVRKENQANTSV 571

Db 181 DLTNLAAPMAVGQIMTFGSPVIGCGFISGMNLSVSMCVYVLLWKVYQKTPALAVKAGLK 240
QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRTFRDGMV 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRTFRDGMV 300
QY 301 SYNOPVFLAGMGLAFLYMTVLGFDCTTGAYATQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNOPVFLAGMGLAFLYMTVLGFDCTTGAYATQGLSGSILSLMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
Db 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
QY 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLFAGVIAARIGLWSFDLTVTQ 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLFAGVIAARIGLWSFDLTVTQ 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
QY 541 FAQNTLGNKLFACGPDKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDKEVRKENQANTSVV 571

RESULT 3
AAB27245
ID AAB27245 standard; Protein; 571 AA.
XX
AC AAB27245;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human EXMAD-23 SEQ ID NO: 23.
XX
EX Extracellular matrix and adhesion-associated protein; EXMAD; cancer;
KW inflammation; reproductive disorder; cardiovascular disorder;
KW immune disorder; musculoskeletal disorder; developmental disorder;
KW gastrointestinal disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200068380-A2.
XX
PD 16-NOV-2000.
XX
PF 10-MAY-2000; 2000WO-US12811.
XX
PR 11-MAY-1999; 99US-0133643.
XX
PR 23-AUG-1999; 99US-0150409.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Bandman O, Hillman JL, Tang YT, Lal P, Yue H, Baughn MR, Lu DAM;
PI Azimzai Y;
XX
XX WPI; 2001-007395/01.
DR N-PSDB; AAC66912.
XX
XX Isolated polynucleotide encoding extracellular matrix or
PT adhesion-associated protein (EXMAD) useful for diagnosing, treating, or
PT preventing disorders associated with expression of EXMAD such as
PT proliferative, immune and genetic disorders -
XX
PS Claim 1; Page 109-110; 129pp; English.
XX
CC The present invention provides the protein and coding sequences for 25
CC novel extracellular matrix and adhesion-associated proteins (EXMADs).
CC These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5,
CC EXMAD-6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12,
CC EXMAD-13, EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19,

CC EXMAD-20, EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are
CC useful in the prevention and treatment of cancers, cell proliferation,
CC cardiovascular, reproductive, immune, musculoskeletal, developmental and
CC gastrointestinal disorders and inflammation.
XX
SQ Sequence 571 AA;
Query Match 100.0%; Score 2929; DB 22; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.7e-271;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTRAGDNNRQRCGCGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFLVELYNSLL 60
Db 1 MTRAGDNNRQRCGCGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFLVELYNSLL 60
QY 61 TAVYGLVAGSVLVGLGIDVDKNARLKVQATSLVVQNVSVILCGIILMMVFLKHHEL 120
Db 61 TAVYGLVAGSVLVGLGIDVDKNARLKVQATSLVVQNVSVILCGIILMMVFLKHHEL 120
QY 121 LTMVHGWLVTSCYLIITIANIANLASTATAITIQRDWIVVVGEDRSKLANMNAITIRRI 180
Db 121 LTMVHGWLVTSCYLIITIANIANLASTATAITIQRDWIVVVGEDRSKLANMNAITIRRI 180
QY 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGMNLSVSMCVYVLLWKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGMNLSVSMCVYVLLWKVYQKTPALAVKAGLK 240
QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRTFRDGMV 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRTFRDGMV 300
QY 301 SYNOPVFLAGMGLAFLYMTVLGFDCTTGAYATQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNOPVFLAGMGLAFLYMTVLGFDCTTGAYATQGLSGSILSLMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
Db 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
QY 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLFAGVIAARIGLWSFDLTVTQ 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLFAGVIAARIGLWSFDLTVTQ 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
QY 541 FAQNTLGNKLFACGPDKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDKEVRKENQANTSVV 571

RESULT 4
AAB23866
ID AAB23866 standard; Protein; 571 AA.
XX
AC AAB23866;
XX
DT 10-SEP-2002 (first entry)
XX
DE Human ferroportin 1 mutant protein.
XX
KW Human; ferroportin 1; hereditary haemochromatosis; therapy; mutant;
KW muteln.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key. Location/Qualifiers
FT Misc-difference 77
FT /note= "Wild type Ala substituted with Asp"
XX
PN WO200233119-A2.

XX 25-APR-2002.
XX 17-OCT-2001; 2001WO-EPI2018.
XX 17-OCT-2000; 2000IT-MI02240.
XX (PIET/) PIETRANGELO A.
XX Pietrangelo A;
XX WPI; 2002-444246/47.
XX N-PSDB; AAD38291.
XX Novel nucleic acid coding for ferroportin 1 useful for the in vitro
PT diagnosis of hereditary hemochromatosis in a mammal, comprises a
XX mutation of the codon coding for the ferroportin 1
XX Claim 13; Page 35-36; 37pp; English.
XX The invention relates to mutations in the gene coding for human
CC ferroportin 1 associated with hereditary hemochromatosis and methods
CC for the diagnosis of hereditary hemochromatosis based on the
CC identification of such mutations. Polynucleotides of the invention
CC are used for the production of recombinant mutated ferroportin 1.
CC its fragment or a chimeric protein including the fragment, in order
CC to study the functional characteristics of the mutated protein. The
CC are also used for diagnostic and therapeutic purposes. The present
CC sequence is human ferroportin 1 mutant protein.
XX SQ Sequence 571 AA;
Query Match 99.9%; Score 2925; DB 23; Length 571;
Best Local Similarity 99.8%; Pred. No. 2.1e-270;
Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTRAGDHNRRGCGSLADYLTSAKFLLYLGHSLSTWGDHMHFAVSFLVLYGNSLL 60
DB 1 MTRAGDHNRRGCGSLADYLTSAKFLLYLGHSLSTWGDHMHFAVSFLVLYGNSLL 60
QY 61 TAVYGLVAGSVLVGLGIIIDGVNDKARLKAQTSLVVQNVSVILCGIILMVFLKHHEL 120
DB 61 TAVYGLVAGSVLVGLGIIIDGVNDKARLKAQTSLVVQNVSVILCGIILMVFLKHHEL 120
QY 121 LTMVHGWLTVSCVLIITIANIANLASTATAITIQRDWIVVAVAGEDRSKLANNNATIRI 180
DB 121 LTMVHGWLTVSCVLIITIANIANLASTATAITIQRDWIVVAVAGEDRSKLANNNATIRI 180
QY 181 DQLTNIIAPWAVGQIMTFGSPVIGCGFISGWNLVSMCVSVLLWKVYQKTPALAVKAGLK 240
DB 181 DQLTNIIAPWAVGQIMTFGSPVIGCGFISGWNLVSMCVSVLLWKVYQKTPALAVKAGLK 240
QY 241 EETELKQLMLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFTRDGVW 300
DB 241 EETELKQLMLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFTRDGVW 300
QY 301 SYYNQPVFLAGMLAFIYMTVGFDCITTCYAYTQGLSGSILILMGASAITGIMGTAVF 360
DB 301 SYYNQPVFLAGMLAFIYMTVGFDCITTCYAYTQGLSGSILILMGASAITGIMGTAVF 360
QY 361 TWLRRKGLVTRTGLISLAQLSCILICVISVFMGSPDLDSVSPFEDIRSRFTQGESITP 420
DB 361 TWLRRKGLVTRTGLISLAQLSCILICVISVFMGSPDLDSVSPFEDIRSRFTQGESITP 420
QY 421 TKIPEITTEIYMGNSNSANIVPETSPEPIISVSLFAGVIAARIGLSMFDLTVTQLL 480
DB 421 TKIPEITTEIYMGNSNSANIVPETSPEPIISVSLFAGVIAARIGLSMFDLTVTQLL 480
QY 481 QENVIESERGIINGVQNSMYNLLDLHLHFIMVILAPNPEAFGLLVLVSFVAMGHIMYFR 540
DB 481 QENVIESERGIINGVQNSMYNLLDLHLHFIMVILAPNPEAFGLLVLVSFVAMGHIMYFR 540
QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

DB 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
RESULT 5
AAV29327
ID AAY29327 standard; Protein; 571 AA.
XX AAY29327;
XX 29-SEP-1999 (first entry)
DE Human secreted protein clone ewl50_1 protein sequence.
XX Human; secreted protein; nutrition; cytokine; cytokine; cell proliferation;
KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
KW tumour invasion suppressor; tumour inhibition.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 202 /label= unknown
FT /note= "encoded by RTC"
FT Misc-difference 504
FT /label= unknown
FT /note= "encoded by RAT"
XX WO9937674-A1.
XX 29-JUL-1999.
XX 21-JAN-1999; 99WO-US01404.
XX 20-JAN-1999; 99US-0235609.
XX 22-JAN-1998; 98US-0072134.
XX (GEMY) GENETICS INST INC.
XX Agostino MJ, Clark HF, Collins-Racie LA, Fecthel K;
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Steininger RJ;
PI Treacy M, Wong GG;
XX WPI; 1999-458682/38.
DR N-PSDB; AAX90442.
XX New polynucleotides encoding secreted human proteins derived from,
PT e.g. fetal brain potentially used as immunostimulators
XX Claim 16; Page 115-117; 139pp; English.
XX The present sequence represents a human secreted protein. Human secreted
CC protein polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful
CC for gene therapy.
XX SQ Sequence 571 AA;
Query Match 99.6%; Score 2917; DB 20; Length 571;
Best Local Similarity 99.6%; Pred. No. 1.2e-269;
Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTRAGDHNRRQCGGSLADYLTSAKFLYLGHSLSTWGDHMHFAVSVFLVELYGNLSLL 60
 Db 1 MTRAGDHNRRQCGGSLADYLTSAKFLYLGHSLSTWGDHMHFAVSVFLVELYGNLSLL 60
 QY 61 TAVYGLVAVGSLVLGAILGDWVDKNAKVAQTSILVQNVSVILCGIILMMVFLHKL 120
 Db 61 TAVYGLVAVGSLVLGAILGDWVDKNAKVAQTSILVQNVSVILCGIILMMVFLHKL 120
 QY 121 LTMVHGWLTSYCIILITIANIANLASTATAITIQDWMIVVAGEDSKLANMATTIRI 180
 Db 121 LTMVHGWLTSYCIILITIANIANLASTATAITIQDWMIVVAGEDSKLANMATTIRI 180
 QY 181 DQTLNLAAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWVYKTPALAVKAGLK 240
 Db 181 DQTLNLAAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWVYKTPALAVKAGLK 240
 QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRFDGW 300
 Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRFDGW 300
 QY 301 SYNQPVFLAGMGLAFLYMTVLGFDCTITGAYTQGLSGSILSILMGASAITGIMTVAF 360
 Db 301 SYNQPVFLAGMGLAFLYMTVLGFDCTITGAYTQGLSGSILSILMGASAITGIMTVAF 360
 QY 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVMPGSPDLVSPEDIRSRFIQGESITP 420
 Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVMPGSPDLVSPEDIRSRFIQGESITP 420
 QY 421 TKIPITTEIYMSNGSNANIPETSPESVPIISVLLPAGVIAIRIGLWGFDTLVTOLL 480
 Db 421 TKIPITTEIYMSNGSNANIPETSPESVPIISVLLPAGVIAIRIGLWGFDTLVTOLL 480
 QY 481 QENVIESERGIINGVQNSMNLDDLHFIWILAPNPEAFGLLVLISVSVFAMGHIMYFR 540
 Db 481 QENVIESERGIINGVQNSMNLDDLHFIWILAPNPEAFGLLVLISVSVFAMGHIMYFR 540
 QY 541 FAQNTLGNKLFACGPDKEVRKENQANTSVV 571
 Db 541 FAQNTLGNKLFACGPDKEVRKENQANTSVV 571

RESULT 6

AAU33053
 ID AAU33053 standard; Protein; 571 AA.

AAU33053;

16-JAN-2002 (first entry)

Human secreted protein ew150_1.

Human; secreted protein; antiinflammatory; immunosuppressive;
 nontropic; neuroprotective; antiarthritis; antimicrobial; vulnary;
 cystostatic; antidiabetic; viricide; antinfertility; anticonvulsant;
 vasotropic; antiparkinsonian; immunostimulant; dermatological;
 antirheumatic; antitumor; antulcer; osteopathic; tranquiliser;
 cerebroprotective; cytokine; cell proliferation; cell differentiation;
 immune deficiency; severe combined immunodeficiency; SCID; tumour;
 autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 Parkinson's disease; Huntington's disease; infection; cardiac disease;
 stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 food supplement; vaccine.

Homo sapiens.

WO200175068-A2.

11-OCT-2001.

22-MAR-2001; 2001WO-US09369.

XX

PR 30-MAR-2000; 2000US-0539330.
 PR 04-DEC-2000; 2000US-0729674.
 XX (GEMY) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
 PI Clark H, Fichtel K, Merberg D;
 XX WPI; 2001-639363/73.
 DR N-PSDB; AAS59271.
 XX

Secreted human proteins, useful as vaccine for treating various diseases such as autoimmune disorders (e.g. multiple sclerosis), and nervous system disorders (e.g. stroke)

Disclosure; Page 542-544; 619pp; English.

The invention relates to novel human secreted proteins, the nucleic acids encoding them. The protein may exhibit cytokine, cell proliferation or cell differentiation activity or may induce production of other cytokines in certain cell populations and may exhibit immune stimulating or immune suppressing activity, which is useful for the treatment of various immune deficiencies and disorders e.g. severe combined immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation. The proteins are also useful in the treatment of diseases and disorders including tissue, skin and organ transplantation and in graft-versus-host diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid cell deficiencies, wound healing and tissue repair, in the treatment of burns, incisions and ulcers; as well as in treatment of periodontal disease, osteoporosis or osteoarthritis, mediated by inflammatory processes, diseases of the peripheral nervous system, Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and central nervous system vessel e.g. stroke, sepsis, inflammatory bowel disease, ulcers, bone regeneration. The protein, having activin- or inhibin-related activities is useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. The proteins and nucleic acids are also useful as food supplements. The present sequence represents a secreted protein of the invention.

Sequence 571 AA;

Query Match 99.6%; Score 2917; DB 22; Length 571;
 Best Local Similarity 99.6%; Pred. No. 1.2e-269;
 Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTRAGDHNRRQCGGSLADYLTSAKFLYLGHSLSTWGDHMHFAVSVFLVELYGNLSLL 60
 Db 1 MTRAGDHNRRQCGGSLADYLTSAKFLYLGHSLSTWGDHMHFAVSVFLVELYGNLSLL 60
 QY 61 TAVYGLVAVGSLVLGAILGDWVDKNAKVAQTSILVQNVSVILCGIILMMVFLHKL 120
 Db 61 TAVYGLVAVGSLVLGAILGDWVDKNAKVAQTSILVQNVSVILCGIILMMVFLHKL 120
 QY 121 LTMVHGWLTSYCIILITIANIANLASTATAITIQDWMIVVAGEDSKLANMATTIRI 180
 Db 121 LTMVHGWLTSYCIILITIANIANLASTATAITIQDWMIVVAGEDSKLANMATTIRI 180
 QY 181 DQTLNLAAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWVYKTPALAVKAGLK 240
 Db 181 DQTLNLAAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWVYKTPALAVKAGLK 240

QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRFDGW 300
 Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRFDGW 300
 QY 301 SYNQPVFLAGMGLAFLYMTVLGFDCTITGAYTQGLSGSILSILMGASAITGIMTVAF 360
 Db 301 SYNQPVFLAGMGLAFLYMTVLGFDCTITGAYTQGLSGSILSILMGASAITGIMTVAF 360

QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFEDIRSRFIQGESITP 420
Db |||||
QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFEDIRSRFIQGESITP 420
Db |||||
QY 421 TKIPEITTEIYMGNGNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Db |||||
QY 481 QENVIESERGIINGVQNMNLLDLHLFIMVILAPNPEAFGLLVLSVSFVANGHIMYFR 540
Db |||||
QY 541 FAQNTLGNKLPACGPDAKEVRKENQANTSVV 571
Db |||||
ID ABB55762 standard; Protein; 571 AA.
AC ABB55762;
XX
DT 14-FEB-2002 (first entry)
XX Human polypeptide SEQ ID NO 130.
XX Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
OS Homo sapiens.
XX
FN US2001039335-A1.
XX
PD 08-NOV-2001.
XX
PF 04-DEC-2000; 2000US-0729674.
XX
PR 10-APR-1997; 97US-126435P.
PR 04-DEC-1997; 97US-067454P.
PR 20-DEC-1997; 97US-068379P.
PR 02-JAN-1998; 98US-070346P.
PR 07-JAN-1998; 98US-070643P.
PR 08-JAN-1998; 98US-070755P.
PR 13-JAN-1998; 98US-071304P.
PR 22-JAN-1998; 98US-072134P.
PR 30-JAN-1998; 98US-073095P.
PR 18-FEB-1998; 98US-075038P.
PR 30-MAR-2000; 2000US-0539330.
PR 23-NOV-1998; 98US-0197886.
XX
PA (JACO/) JACOBS K.
PA (MCCO/) MCCOY J M.
PA (LAVA/) LAVALLIE E R.
PA (COLL/) COLLINS-RACIE L A.
PA (EVAN/) EVANS C.
PA (MERB/) MERBERG D.
PA (TREA/) TREACY M.
PA (AGOS/) AGOSTINO M J.
PA (STEI/) STEININGER R J.
PA (SPAU/) SPAULDING V.
PA (WONG/) WONG G G.
PA (CLAR/) CLARK H.
PA (FECH/) FECHTEL K.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
PI Wong GG, Clark H, Fechtel K;
XX WPI; 2002-040725/05.
DR N-PSDB; ABA90940.
XX
PT New secreted proteins and encoding polynucleotides, useful in gene
PT therapies, particularly for preventing or treating autoimmune
PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
PT stroke or inflammations
XX
PS Disclosure; Page 270-271; 349pp; English.
XX
CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
CC ABA90980) and encoded proteins (ABBS5698-ABBS5800), especially
CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
CC proteins SEQ ID NO 2 (ABBS5698) and SEQ ID NO 20 (ABBS5707) contained in
CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
CC are deposited with the American Type Culture Collection (ATCC) with
CC accession number 98599. The polynucleotides and encoded polypeptides have
CC cytostatic, anti-inflammatory, immunomodulator, vulnerary,
CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
CC and anti-inflammatory activity and acting as cytokine modulators,
CC haematopoiesis regulators, tissue growth modulators and/or cadherin
CC suppressors. The polypeptides and polynucleotides are useful in gene
CC therapies, particularly for preventing, treating or ameliorating any of
CC the following diseases: immune deficiency and disorders; e.g. bacterial
CC or fungal infections, autoimmune disorders, cancer, systemic lupus
CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
CC osteoarthritis; central and peripheral nervous system diseases and
CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis or Shy-brager syndrome;
CC haemophilia, cardiac infarction or stroke, inflammations, shock, sepsis
CC or systemic inflammatory response syndrome, ischaemia-reperfusion
CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or
CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
CC foliaceus.
XX
SQ Sequence 571 AA;
Query Match 99.6%; Score 2917; DB 23; Length 571;
Best Local Similarity 99.6%; Pred. No. 1.2e-269;
Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MTRAGDHNRRQCCGSLADYLTSAKFLYLGHSLSTWGRMMHFAVSFLVELYNSLLL 60
Db |||||
QY 1 MTRAGDHNRRQCCGSLADYLTSAKFLYLGHSLSTWGRMMHFAVSFLVELYNSLLL 60
Db |||||
QY 61 TAVYGLVWAGSVLVLGAIGDWDVKNARLKVQAOTSLVQNVSVILCGIILMMVFLHKHEL 120
Db |||||
QY 61 TAVYGLVWAGSVLVLGAIGDWDVKNARLKVQAOTSLVQNVSVILCGIILMMVFLHKHEL 120
Db |||||
QY 121 LTMVHGWLTSYIIIIITIANIANLASTATAITIQDNIWVVGEDRSKLANNNATIRI 180
Db |||||
QY 121 LTMVHGWLTSYIIIIITIANIANLASTATAITIQDNIWVVGEDRSKLANNNATIRI 180
Db |||||
QY 181 DOLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSNKVEYVLLWKVYQKTPALAVKAGLK 240
Db |||||
QY 181 DOLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSNKVEYVLLWKVYQKTPALAVKAGLK 240
Db |||||
QY 241 BEETELKQNLNHLKDEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGVV 300
Db |||||
QY 241 BEETELKQNLNHLKDEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGVV 300
Db |||||
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTITGAYTQGLSGSILSLMGASAITGIMGTAF 360
Db |||||
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTITGAYTQGLSGSILSLMGASAITGIMGTAF 360
Db |||||
QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFEDIRSRFIQGESITP 420
Db |||||
QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFEDIRSRFIQGESITP 420
Db |||||

QY 421 TKIPETITTYMNSGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPETITTYMNSGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQLL 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFWILAPNPEAFGLLVLSVSFVAMGHIMYPR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHLFWILAPNPEAFGLLVLSVSFVAMGHIMYPR 540
QY 541 FRAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FRAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

RESULT 8
ABP41875
ID ABP41875 standard; Protein; 408 AA.
XX AC ABP41875;
XX DT 22-AUG-2002 (first entry)
XX DE Human ovarian antigen HISBE12, SEQ ID NO:3007.
XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX OS Homo sapiens.
XX WO200200677-A1.
XX PD 03-JAN-2002.
XX PF 07-JUN-2001; 2001WO-US18569.
XX PR 07-JUN-2000; 2000US-209467P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
XX DR N-PSDB; ABQ54952.
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX PS Claim 11; SEQ ID NO 3007; 2922pp; English.
XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 408 AA;

Query Match 68.9%; Score 2018; DB 23; Length 408;
Best Local Similarity 100.0%; Pred. No. 6.6e-184;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 RIDQLTNILAPMAVGOIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKYQKTPALAVKAG 238
Db 16 RIDQLTNILAPMAVGOIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKYQKTPALAVKAG 75
QY 239 LKEEETELKQNLNKHDKTEPKLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPPRTFRDG 298
Db 76 LKEEETELKQNLNKHDKTEPKLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPPRTFRDG 135
QY 299 WVSYYNPVFLAGMGLAFYMTVLGFCITTYAYTQGLSGSILSLMGASAITGIMGTV 358
Db 136 WVSYYNPVFLAGMGLAFYMTVLGFCITTYAYTQGLSGSILSLMGASAITGIMGTV 195
QY 359 AFTWLRKCGLVRTLISGLAQLSCLILCVISVMPGSPDLVSPPEDISRFTQGESI 418
Db 196 AFTWLRKCGLVRTLISGLAQLSCLILCVISVMPGSPDLVSPPEDISRFTQGESI 255
QY 419 TPTKIPETITTYMNSGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQ 478
Db 256 TPTKIPETITTYMNSGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQ 315
QY 479 LLQENVIESERGIINGVQNSMNYLLDLHLFWILAPNPEAFGLLVLSVSFVAMGHIMY 538
Db 316 LLQENVIESERGIINGVQNSMNYLLDLHLFWILAPNPEAFGLLVLSVSFVAMGHIMY 375
QY 539 FRAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 376 FRAQNTLGNKLFACGPDPAKEVRKENQANTSVV 408

RESULT 9

ABB89115
ID ABB89115 standard; Protein; 382 AA.

XX AC ABB89115;

XX DT 24-MAY-2002 (first entry)

XX DE Human polypeptide SEQ ID NO 1491.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

XX OS Homo sapiens.

XX WO200190304-A2.

XX PD 29-NOV-2001.

XX PF 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI; 2002-122018/16.
DR N-PSDB; ABL9524.
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX Claim 11; SEQ ID NO 1491; 2081pp + Sequence Listing; English.
XX The invention relates to novel genes (ABL9449-ABL90853) and proteins
CC (ABB85040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 382 AA;
Query Match 61.1%; Score 1789; DB 23; Length 382;
Best Local Similarity 92.4%; Pred. No. 4.7e-162;
Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;
QY 190 MAVGQIMTFSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEBETELKQL 249
DB 1 MAVGQIMTFSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEBETELKQL 60
QY 250 NLHKDTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTRDGMVSYNQPVFL 309
DB 61 NLHKDTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTRDGMVSYNQPVFL 120
QY 310 AGMGLAFLYMTVLGFDCTTTCYAYTQGLSGSILSILMGASAITGIMGTVAFTWLRKCGL 369
DB 121 AGMGLAFLYMTVLGFDCTTTCYAYTQGLSGSILSILMGASAITGIMGTVAFTWLRKCGL 180
QY 370 VRTGLISGLAQLCLILCVISVFMPSGPDLSVSPEDIRSFRTQGESITPKIPETTE 429
DB 181 GSAGLISGLAQLCLILCVISVFMPSGPDLSVSPEDIRSFRTQGESITPKIPETTE 240
QY 430 IYMSNGSNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQLLQENVIESER 489
DB 241 IYMSNGSNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQLLQENVIESER 300
QY 490 GIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFRFAQNTLGNK 549
DB 301 GIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFRFAQNTLGNK 360
QY 550 LFACGPDPAKVRKENQANTSV 571
DB 361 LFACGPDPAKVRKENQANTSV 382
RESULT 10
AAW74869

ID XX
AC XX
XX XX
DT XX
XX XX
DE XX
XX XX
XX XX
KW Human; secreted protein; testis; tumour; foetal brain tissue;
KW fusion protein; cancer; central nervous system; seizure;
KW diagnosis; neurodegenerative disease.
XX OS
XX Homo sapiens.
XX Key
PH Misc-difference 383
FT /label= unknown
XX
XX WO9839448-A2.
PN 11-SEP-1998.
PD 06-MAR-1998; 98WO-US04493.
XX 02-OCT-1997; 97US-0061060.
XX 07-MAR-1997; 97US-0038621.
XX 07-MAR-1997; 97US-0040161.
XX 07-MAR-1997; 97US-0040162.
XX 07-MAR-1997; 97US-0040163.
XX 07-MAR-1997; 97US-0040333.
XX 07-MAR-1997; 97US-0040334.
XX 07-MAR-1997; 97US-0040336.
XX 07-MAR-1997; 97US-0040626.
XX 11-APR-1997; 97US-0043311.
XX 11-APR-1997; 97US-0043312.
XX 11-APR-1997; 97US-0043313.
XX 11-APR-1997; 97US-0043314.
XX 11-APR-1997; 97US-0043568.
XX 11-APR-1997; 97US-0043569.
XX 11-APR-1997; 97US-0043576.
XX 11-APR-1997; 97US-0043578.
XX 11-APR-1997; 97US-0043580.
XX 11-APR-1997; 97US-0043669.
XX 11-APR-1997; 97US-0043670.
XX 11-APR-1997; 97US-0043671.
XX 11-APR-1997; 97US-0043672.
XX 11-APR-1997; 97US-0043674.
XX 23-MAY-1997; 97US-0047492.
XX 23-MAY-1997; 97US-0047500.
XX 23-MAY-1997; 97US-0047501.
XX 23-MAY-1997; 97US-0047502.
XX 23-MAY-1997; 97US-0047503.
XX 23-MAY-1997; 97US-0047581.
XX 23-MAY-1997; 97US-0047582.
XX 23-MAY-1997; 97US-0047583.
XX 23-MAY-1997; 97US-0047584.
XX 23-MAY-1997; 97US-0047585.
XX 23-MAY-1997; 97US-0047586.
XX 23-MAY-1997; 97US-0047587.
XX 23-MAY-1997; 97US-0047588.
XX 23-MAY-1997; 97US-0047589.
XX 23-MAY-1997; 97US-0047590.
XX 23-MAY-1997; 97US-0047592.
XX 23-MAY-1997; 97US-0047593.
XX 23-MAY-1997; 97US-0047594.
XX 23-MAY-1997; 97US-0047595.
XX 23-MAY-1997; 97US-0047596.
XX 23-MAY-1997; 97US-0047597.
XX 23-MAY-1997; 97US-0047598.
XX 23-MAY-1997; 97US-0047599.
XX 23-MAY-1997; 97US-0047600.
XX 23-MAY-1997; 97US-0047601.
XX 23-MAY-1997; 97US-0047612.

PR	07-MAR-1997;	97US-040161P.	PR	22-AUG-1997;	97US-056875P.	XX	(HUMA-) HUMAN GENOME SCI INC.
PR	07-MAR-1997;	97US-040162P.	PR	22-AUG-1997;	97US-056876P.	XX	
PR	07-MAR-1997;	97US-040163P.	PR	22-AUG-1997;	97US-056877P.	PI	Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC, Bednarik DR;
PR	07-MAR-1997;	97US-040333P.	PR	22-AUG-1997;	97US-056878P.	PI	Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrie AM;
PR	07-MAR-1997;	97US-040334P.	PR	22-AUG-1997;	97US-056879P.	PI	Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;
PR	07-MAR-1997;	97US-040336P.	PR	22-AUG-1997;	97US-056880P.	PI	Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
PR	07-MAR-1997;	97US-040626P.	PR	22-AUG-1997;	97US-056881P.	XX	
PR	11-APR-1997;	97US-043311P.	PR	22-AUG-1997;	97US-056882P.	DR	WPI; 2002-634796/68.
PR	11-APR-1997;	97US-043313P.	PR	22-AUG-1997;	97US-056884P.	DR	N-PSDB; ABS73638.
PR	11-APR-1997;	97US-043314P.	PR	22-AUG-1997;	97US-056886P.	XX	
PR	11-APR-1997;	97US-043315P.	PR	22-AUG-1997;	97US-056887P.	XX	
PR	11-APR-1997;	97US-043368P.	PR	22-AUG-1997;	97US-056889P.	PT	New isolated human secreted protein for diagnosing, preventing,
PR	11-APR-1997;	97US-043569P.	PR	22-AUG-1997;	97US-056892P.	PT	treating or ameliorating medical conditions and used as a food additive
PR	11-APR-1997;	97US-043576P.	PR	22-AUG-1997;	97US-056893P.	PT	or preservative -
PR	11-APR-1997;	97US-043578P.	PR	22-AUG-1997;	97US-056894P.	PS	Example 1; SEQ ID No 460; 129pp; English.
PR	11-APR-1997;	97US-043580P.	PR	22-AUG-1997;	97US-056903P.	CC	The invention relates to an isolated protein that is one of 186 human
PR	11-APR-1997;	97US-043586P.	PR	22-AUG-1997;	97US-0435869P.	CC	secreted proteins, given in the specification, encoded by one of
PR	11-APR-1997;	97US-043670P.	PR	22-AUG-1997;	97US-043671P.	CC	309 cDNA sequences also given in the specification. The protein is used
PR	11-APR-1997;	97US-043671P.	PR	22-AUG-1997;	97US-043672P.	CC	in a pharmaceutical composition used to prevent, treat or ameliorate a
PR	11-APR-1997;	97US-043674P.	PR	22-AUG-1997;	97US-043674P.	CC	medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
PR	11-APR-1997;	97US-043672P.	PR	22-AUG-1997;	97US-043674P.	CC	dogs, chickens or sheep. Disorders which are diagnosed or treated include
PR	23-MAY-1997;	97US-047492P.	PR	05-SEP-1997;	97US-057761P.	CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
PR	23-MAY-1997;	97US-047500P.	PR	05-SEP-1997;	97US-057650P.	CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
PR	23-MAY-1997;	97US-047501P.	PR	05-SEP-1997;	97US-057669P.	CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
PR	23-MAY-1997;	97US-047502P.	PR	12-SEP-1997;	97US-058785P.	CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease,
PR	23-MAY-1997;	97US-047581P.	PR	02-OCT-1997;	97US-061060P.	CC	infections caused by bacteria, viruses and fungi and ocular disorders
PR	23-MAY-1997;	97US-047582P.	XX			CC	e.g. corneal infection. The polypeptides can also be used to aid wound
PR	23-MAY-1997;	97US-047583P.	XX			CC	healing and epithelial cell proliferation, to prevent skin aging due to
PR	23-MAY-1997;	97US-047584P.	XX			CC	sunburn, to maintain organs before transplantation, for supporting cell
PR	23-MAY-1997;	97US-047585P.	XX			CC	culture of primary tissues, to regenerate tissues and in chemotaxis. The
PR	23-MAY-1997;	97US-047586P.	XX			CC	polypeptides can also be used as a food additive or preservative to
PR	23-MAY-1997;	97US-047587P.	XX			CC	increase or decrease storage capabilities, fat content, lipid, protein,
PR	23-MAY-1997;	97US-047588P.	XX			CC	carbohydrate, vitamins, minerals, cofactors and other nutritional
PR	23-MAY-1997;	97US-047589P.	XX			CC	components. The present sequence represents one of the novel human
PR	23-MAY-1997;	97US-047590P.	XX			CC	secreted proteins of the invention.
PR	23-MAY-1997;	97US-047591P.	XX			CC	Note: This sequence did not form part of the printed specification,
PR	23-MAY-1997;	97US-047592P.	XX			CC	but was obtained in electronic format directly from USPTO at
PR	23-MAY-1997;	97US-047593P.	XX			CC	seqdata.uspto.gov/sequence.html?DocID=6420526B1.
PR	23-MAY-1997;	97US-047594P.	XX			XX	
PR	23-MAY-1997;	97US-047595P.	XX			XX	
PR	23-MAY-1997;	97US-047596P.	XX			XX	
PR	23-MAY-1997;	97US-047597P.	XX			XX	
PR	23-MAY-1997;	97US-047598P.	XX			XX	
PR	23-MAY-1997;	97US-047599P.	XX			XX	
PR	23-MAY-1997;	97US-047600P.	XX			XX	
PR	23						

QY 190 MAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLKEEETELKQL 249
Db 1 MAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLKEEETELKQL 60
QY 250 NLHKDTEPKPLEGTHLMGVKDSNIHELEHQBEPICASQMAEPFRFRDQGWVSYNQPVFL 309
Db 61 NLHKDTEPKPLEGTHLMGVKDSNIHELEHQBEPICASQMAEPFRFRDQGWVSYNQPVFL 120
QY 310 AGMGLAFYMTVLGFDCTTGVAYTQGLSGSILSLMGASAITGIMGTVAFTLRRKGL 369
Db 121 AGMGLAFYMTVLGFDCTTGVAYTQGLSGSILSLMGASAITGIMGTVAFTLRRKGL 180
QY 370 VRTGLISGLAQSLCILCVISVFMPSGLDLSVPFEDIRSRFQGESITPKIPEITTE 429
Db 181 GSAGLISGLAQSLCILCVISVFMPSGLDLSVPFEDIRSRFQGESITPKIPEITTE 240
QY 430 IYMSGNSANIVPETSPEPIISVLLFAGVIAARIGLSFDTVTQLLQENVIESER 489
Db 241 IYMSGNSANIVPETSPEPIISVLLFAGVIAARIGLSFDTVTQLLQENVIESER 300
QY 490 GIINGVQSNMNYLLDLHFIMVILAPNEAFGLLVLSVFMVAMCHIMYFPAQNTLGNK 549
Db 301 GIINGVQSNMNYLLDLHFIMVILAPNEAFGLLVLSVFMVAMCHIMYFPAQNTLGNK 360
QY 550 LFACGPDAAKEVRKENQANTSVV 571
Db 361 LFACGPDAAKEVRKENQANTSVV 382

RESULT 12
AAM40920

ID AAM40920 standard; Protein; 439 AA.

AC AAM40920;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Example 2; SEQ ID NO 5851; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
SQ Sequence 439 AA;

Query Match 55.7%; Score 1630; DB 22; Length 439;

Best Local Similarity 78.9%; Pred. No. 9.3e-147;

Matches 333; Conservative 8; Mismatches 31; Indels 50; Gaps 6;

QY 1 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGRMWHFAVSFVLVYNSLL 60
Db 35 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGRMWHFAVSFVLVYNSLL 94
QY 61 TAVYGLVWAGSVLVLGAILGWDVKNARLKVQOTSLVQNVSVLTCGIIIMVFLHKL 120
Db 95 TAVYGLVWAGSVLVLGAILGWDVKNARLKVQOTSLVQNVSVLTCGIIIMVFLHKL 154
QY 121 LTMYHGWLTSCYILITITIANIANLASTATAITQRDWIVVWAGEDRSKLANNNATIRI 180
Db 155 LTMYHGWLTSCYILITITIANIANLASTATAITQRDWIVVWAGEDRSKLANNNATIRI 214
QY 181 DQLTNLAPMAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLK 240
Db 215 DQLTNLAPMAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLK 274
QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHQBEPICASQMAEPFRFRDQGW 300
Db 275 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHQBEPICASQMAEPFRFRDQGW 334
QY 301 SYYNQPVFLAGMGLAF-LYMTVLGFDG-----ITTGAYTQGLSGSILSILM 346
Db 335 SYYNQPVFLGWHGSCFPLY-----DCPGLXLLHHHVRVRLHSGTEWFFHPQYFDSI----- 383
QY 347 GASAITGIMGTVAFTWLRRK-----GLVRTGLISGLAQLSCLI--LCVISVFM 393
Db 384 -----SYNNWNGNSPYLATSQMWFGSDRLRIGTAFGLFDVLCDCIHAWK 431
QY 394 PG 395
Db 432 PG 433

RESULT 13

ABG20262

ID ABG20262 standard; Protein; 439 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Novel human diagnostic protein #20253.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:30:57 ; Search time 23 Seconds
(without alignments)
1050.413 Million cell updates/sec

Title: US-09-715-927-6
Perfect score: 2929
Sequence: 1 MTRAGDHRQRCCGSLADY.....ACGPDAKEVRKENQANTSVW 571

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	61.1	383	4	US-09-149-476-460
2	123.5	4.2	524	2	US-08-928-632-12
3	123.5	4.2	524	2	US-09-339-972-12
4	112	3.8	305	4	US-09-252-991A-33032
5	110	3.8	429	4	US-09-328-352-4392
6	109.5	3.7	470	4	US-09-134-001C-4610
7	108	3.7	330	1	US-08-118-270-19
8	108	3.7	330	5	PCT-US93-08528-19
9	108	3.7	618	4	US-08-595-553A-2
10	108	3.7	618	4	US-09-640-198D-4
11	107	3.7	411	4	US-09-134-001C-3299
12	107	3.7	465	4	US-09-328-352-5222
13	104	3.6	534	2	US-09-031-332-4
14	104	3.6	534	3	US-09-299-549-4
15	104	3.6	534	4	US-09-610-417-4
16	103.5	3.5	471	4	US-09-328-314-17
17	103	3.5	358	2	US-08-748-485-6
18	102.5	3.5	529	4	US-09-291-922-28
19	102	3.5	553	4	US-09-134-001C-2974
20	101.5	3.5	996	4	US-09-252-991A-27018
21	100.5	3.4	411	3	US-09-236-080-6
22	100.5	3.4	760	4	US-09-252-991A-27790
23	100	3.4	499	4	US-09-134-001C-5370
24	99.5	3.4	461	1	US-08-194-338-4
25	98.5	3.4	411	3	US-09-236-080-2
26	98.5	3.4	411	4	US-09-336-643A-83
27	98	3.3	366	4	US-09-134-001C-5502

28	98	3.3	443	4	US-09-134-001C-3227	Sequence 3227, Ap
29	98	3.3	808	4	US-09-134-001C-3105	Sequence 3105, Ap
30	97.5	3.3	596	4	US-09-252-991A-26031	Sequence 26031, A
31	97.5	3.3	808	2	US-08-629-231A-33	Sequence 33, Appl
32	97.5	3.3	808	2	US-08-658-335B-33	Sequence 33, Appl
33	97.5	3.3	808	4	US-09-406-640-33	Sequence 33, Appl
34	97.5	3.3	1033	4	US-09-252-991A-20611	Sequence 20611, A
35	97	3.3	359	2	US-08-103-170-4	Sequence 4, Appli
36	97	3.3	359	2	US-08-103-170-6	Sequence 6, Appli
37	97	3.3	457	2	US-08-882-704A-6	Sequence 6, Appli
38	97	3.3	457	4	US-09-151-957-6	Sequence 6, Appli
39	97	3.3	484	4	US-09-266-965-17	Sequence 17, Appl
40	96	3.3	427	4	US-09-107-532A-5530	Sequence 5530, Ap
41	96	3.3	1527	4	US-09-376-330-2	Sequence 2, Appli
42	95.5	3.3	370	4	US-09-144-914-8	Sequence 8, Appli
43	95.5	3.3	480	4	US-09-107-532A-6160	Sequence 6160, Ap
44	95	3.2	1421	4	US-09-252-991A-17805	Sequence 17805, A
45	94.5	3.2	398	4	US-09-134-001C-4353	Sequence 4353, Ap

ALIGNMENTS

RESULT 1
US-09-149-476-460
; Sequence 460, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23


```
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match      61.1%; Score 1789; DB 4; Length 383;
Best Local Similarity 92.4%; Pred. No. 8.8e-175;
Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY 190 MAVGQIMTFGSPVTCGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLKEEETELKQL 249
DB 1 MAVGQIMTFGSPVTCGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLKEEETELKQL 60

QY 250 NLHKDEPKLEGLHMGKDSNIHELEHEOEPTCASQMAEPFRFRDGMVSYYNQPVFL 309
DB 61 NLHKDEPKLEGLHMGKDSNIHELEHEOEPTCASQMAEPFRFRDGMVSYYNQPVFL 120

QY 310 AGMGLAFIYMTVGFDCITTYAYTQGLSGSILSMGASAITGIMGTVAFTWLRKCGL 369
DB 121 AGMGLAFIYMTVGFDCITTYAYTQGLSGFHPQVFDGSIYNNWNGCSFVLATSKMWF 180

QY 370 VRTGLISGLAQLSCLILCVISVFMPCGSPDLDSVSPFEDIRSRFOGESITTKIPEITTE 429
DB 181 GSAGLISGLAQLSCLILCVISVFMPCGSPDLDSVSPFEDIRSRFOGESITTKIPEITTE 240

QY 430 IYMSNGSNSANIVPTSPESPIISVSLFAGVIAARIGLWSFDLTVTQLLOENVIESER 489
DB 241 IYMSNGSNSANIVPTSPESPIISVSLFAGVIAARIGLWSFDLTVTQLLOENVIESER 300

QY 490 GIINGVQSNMYLLDLHLFIMVILAPNEAFGLLVLSVSVFAMGHIMYFRFAQNTLGNK 549
DB 301 GIINGVQSNMYLLDLHLFIMVILAPNEAFGLLVLSVSVFAMGHIMYFRFAQNTLGNK 360

QY 550 LFACGPDAKEVRKENQANTSVV 571
DB 361 LFACGPDAKEVRKENQANTSVV 382

RESULT 2
US-08-928-692-12
; Sequence 12, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958727 No. 5958727disk of No. 5958727th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944,200-US
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5958727e
US-08-928-692-12

Query Match      4.2%; Score 123.5; DB 2; Length 524;
Best Local Similarity 20.3%; Pred. No. 0.00084;
Matches 99; Conservative 69; Mismatches 150; Indels 169; Gaps 23;

QY 41 MWFHFAVSVELV-----ELYGNSL--LITAVYGLVAGSVLVGLGAIIGDWVDKNARLKVAQ 93
DB 96 LMSLSVSFAVGVMTASFFGWLGTGRKAMLVANILSLVGALLMGF-----SKLGP 149

QY 94 TSLVV--QNVSVILCGIILMVLHKLHETVYHGWLTSCYILIIITIANIANLASTAT 150
DB 150 SHILIIAGRSISGLYCGLI-----SGLVPMY-----IGEIAPTA- 183

QY 151 AITIQDWMVVVAGEDRSKLANMNATRRIDQLNLPMAVGQIMTFGSPVCGFTSG 210
DB 184 -----LARGALGTFHQLAIVTGILI-----SQIIGLEFILG 213

QY 211 -----WNL-----VSMCEVYVLLMKVYQKTPALAVKAGLKEEETELKQLNLHKDEPKPL 260
DB 214 NYDLWHILLGLSGVRAIRLQSLLLFFCPSPRYLYIKL---DEEVKAKQ-SLKR----- 262

QY 261 EGTLMGWKD--SNIHELEHEOEPTCASQMAEPFRFRDGMVSYYNQPVFLAGM-GLAFL 317
DB 263 ----LRGYDDVTQKDNEMRKEREASSEQKVIILFTN---SSYRQDILVALMLHVAQQ 315

QY 318 YMTVLGFDCTTYAYTQGLSGSILSMGASAITGIMGTVAFTWLRKCGLVRTGLISG 377
DB 316 FSGINGIFYYSTISFTAGISKPVYATI-GVGAVNMVF-TAVSVFLVEKAGRRSLFLIGM 373

QY 378 LAQLSCLL-----LCVISVFMPCGSPDLDSVSPFEDIRSRFOGESITP 420
DB 374 SGMFVCAIFMSVGLVLLNKFMSMSYVSMIAIF-----LFVSFFE-----IGP 415

QY 421 TKIPEITTEIYMSNGSNSANIVPTSPESPIIS-----PRPALAIAAFSNWTCNFVALCFQYIADFCGPYVFFL 458
DB 416 GPIPFMVAEFFSQG-----PRPALAIAAFSNWTCNFVALCFQYIADFCGPYVFFL 468

QY 459 FAGVIAA 465
DB 469 FAGVLLA 475

RESULT 3
US-09-339-372-12
; Sequence 12, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6323002 No. 6323002disk of No. 6323002th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```


Db 4 YIDIRYFYNDLSHYNFLLTMTKTSHQSFOSLYM-LIGSAIILALSLGRVHGFGLYL 62
Qy 63 -----VGLVAGSVLVLCAL-IGDWDVKARLKVAQTSLWQVNSVILCG 107
Db 63 VPMSEFGWGQHVFLSAMQNLWGAQOPFGAIDKYG-----SKVVTVGGLLYT 115
Qy 108 I-ILMMVFLKHLLTMTWGHVLTSCYILITIANIANLASTATAITORDWVWVAGED 166
Db 116 LGLLLMAF-----SSSVLILNLSGLIIGLALSATSFTVLLSAVGRAAPPE 161
Qy 167 RSKLANMATIRRIDOLNILAPMAVGQIMFGSPVIGCGFISGWN-----LVSMCV 218
Db 162 KRSMAWGAS-----AAGSFGQFIMLPSTLLLLKTV-GWSSALMVSALLIALII 209
Qy 219 EYVLLWVYQKTPALAVKAGLKEETELKQNLHMDTEPKP-----LEG 262
Db 210 P--LAWML--KGPSNQTPKAIAQPOLTFKQV-LHARKHKEFWLALGFLVCGQVFLG 264
Qy 263 THLMGVKDSNIHELEHEOEPCTAS-----QMAEPFRTRGQW--SYYNQPVFLAGM--- 312
Db 265 VHLPG-----YLIDHGFDAITGTVFLALVGLFNIVGTGAGWLGDRFSKPKLLMALYGS 318
Qy 313 -GLAFLYMTVLGFDCTITGYAY-----TQGLSGSILSILMGASAITGIMGTVAFT 361
Db 319 RGIIAIAFLPLSTYTV-YAFGIIMGLLWSTVPLTNGIVANMGVVKYLSMLSGIVEFT 377
Qy 362 WLRKCGLVRTGLISGL-----AQLSCLILCVISVFMFG 395
Db 378 ---HQVSGFGGWLGVNHDLTGYNVNLWLSIALSIIG 413

RESULT 6

US-09-134-001C-4610
; Sequence 4610, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4610
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4610

Query Match 3.7%; Score 109.5; DB 4; Length 470;
Best Local Similarity 19.5%; Pred. No. 0.019;
Matches 100; Conservative 85; Mismatches 190; Indels 137; Gaps 21;

Qy 73 LVLAGAIGDWD-----KNARKLVAQTSLVQVNSVILCGIILMMVFL 115
Db 45 LILGTFIDAYDITLGMTDQLTQQPHLSPATISVNTSLPIGALFGALLGTLGYQFG 104
Qy 116 HKHELLTMYHGWLTSCYILITIANI-ANLASTATAITIORDWIVVWAGEDR----- 167
Db 105 RKR-----ILSIALLLTWTSLGAALPANNVILICRCIMGFAIGNDSPVATFI 154
Qy 168 SKLANMATIRRID--QNTNIPAMVGQIMTFSPVIGCG-----FTSGNWLVMCMVEY 220
Db 155 AEISNLKHGRNVNVQVWVVAIVTSALVVT-AFFMLGAGHLWRVAVGFGALIAFVLY 213
Qy 221 VLLMKVYQKTPALAVKAGLKEEETELKQNLHMDTEPKPLEGTHLMGVKDSNIHELEHQ 280
Db 214 ILRIKYLHESPTWVNHYSLEKATEFIRKYYHKD-----IHLEGTLEDLNL----- 259

Qy 281 EPTCASQMAEPFRTRFRDQWVSYYNQPVFLA-----GMGLAFLYMTVLGFDCT 328
Db 260 -----SDVTSHPNSWTDLFPKRYIKRIILATAISTALQGMQYVGVG---LYPIIA----- 306
Qy 329 TGYAYTOGLSSILSILMGASA--ITGIMGTVAFTWLRRKGLVTRTGLISGLAQLSCLIL 386
Db 307 -----TVLISKDKTGVLLGTAVNIAIGTLGAYLAQLTLYKLGRKLTMI-GFTLVLLSMV 360
Qy 387 CVISVFMFGSPDLVSVPFEDIRSRFIOGESITP-TKPIETITTEIYMSNGSNSANIVPET 445
Db 361 CV-GLFVHLLPMLNTP-----LIGLFLHSGGGGTQGTQKTICALSFPHLSQATGFVES 415
Qy 446 SPESVPIISVLLFAGVIAARIIGLWSDLTVTQLLOENVIESERGIINGVQNSMNYLLDL 505
Db 416 VSRTSIIIGTFVF--PIILAAVGLTN----- 439
Qy 506 LHFTMVLAPNPEAFGLLVLSVSFVAMG-HI 536
Db 440 ---TMLILSIVP-LLGIIITTSIKWEAVGRKH 467

RESULT 7

US-08-118-270-19
; Sequence 19, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2A
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-118-270-19

Query Match 3.7%; Score 108; DB 1; Length 330;
Best Local Similarity 24.4%; Pred. No. 0.016;
Matches 51; Conservative 39; Mismatches 71; Indels 48; Gaps 11;

Qy 60 LTAVYGLV-----VAGSVLVLAGIIGDWDKNARKLVAQTSLVQVNSVILCGIILMMVFL 115
Db 4 LAAVVGFLIVTVVGVNVLVIAVL-----TSRALRAPQNLFLVSIASADILVATLVMPFS 58

```

Qy 116 HKHELLTMVHGWLTVSCYILITI---ANIANLASTATAITIQRDWLVVVVAGDRSKLAN 172
Db 59 LANEIMWYFGQWGVGLAIDLVLCTSSIVHLC---AISLDRYMSVTOAVE-----YN 109
Qy 173 MNATRRIDQLTNILAPMAVGQIMTF-----GSPVIGCGF-ISGNWLVSMCVB- 219
Db 110 LKRTPRVK---ATIVAVWLISAVISFPPLSVLYRQPDGAAYPQGLNDETWYILSSCIGS 167
Qy 220 -----YVLLW-KVY-----QKTPALAVK 236
Db 168 FFAPCLIYLLVYARIYRVAKRTRTLSEK 196

RESULT 8
PCT-US93-08528-19
; Sequence 19, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-19

Query Match 3.7%; Score 108; DB 5; Length 330;
Best Local Similarity 24.4%; Pred. NO. 0.016;
Matches 51; Conservative 39; Mismatches 71; Indels 48; Gaps 11

Qy 60 LTAIVGLV-----VAGSVLVLGAIGDWVKNARKVAOTSLVYQNVSVILCGIILMMVPL 115
Db 4 LAAVVGFLIVFTVGVGNLVIVTAVL-----TSRAURAFQNLFLVSIASADILVATLVWVPS 58
Qy 116 HKHELLTMVHGWLTVSCYILITI---ANIANLASTATAITIQRDWLVVVVAGDRSKLAN 172
Db 59 LANEIMWYFGQWGVGLAIDLVLCTSSIVHLC---AISLDRYMSVTOAVE-----YN 109
Qy 173 MNATRRIDQLTNILAPMAVGQIMTF-----GSPVIGCGF-ISGNWLVSMCVB- 219
Db 110 LKRTPRVK---ATIVAVWLISAVISFPPLSVLYRQPDGAAYPQGLNDETWYILSSCIGS 167

```

Db 301 --INVVYKCDPILLTGRISAPDQYMPDLLVLDIFEDLPGVGLFLACAYSGTILSTASTS 358
Qy 265 LMVGRKDSNIHELEHQEPTCASQMAEPRTFRDGNVSYNPOVFVLAGMGLAFLYMTVLGF 324
Db 359 INAMAATVEDLIKPRMPGLAPR-----KLVFIS-KGLSFIY-----GS 396
Qy 325 DCITTYAYTQGLSGSILMGLASAITG---IMGTV-----AFT--WLRRKQGL--VR 371
Db 397 ACLTV-----AALSSLLGGVLOGSFTVMGVISGPLLGAFTLGMLLPACNTPGVL 446
Qy 372 TGLISGLAQLSCLILCVISVMPGSPDLDSVPEDIRSRFIQGESITPTKIPBITTEIY 431
Db 447 SGLAAGLA-VSLWAVGATLYPPGE-----QTMGVLP TSAAGCTNDSV 488
Qy 432 M---SNGSNANIVPETSPE-----SVPIISVLLFAGVIAARIGLWSPDLTWTQLL 480
Db 489 LLGPPGATNASGIPSSGMDTGRPALADTFVAISLYLYGALG-----TLTMTL 536
RESULT 10
US-09-640-198D-4
; Sequence 4, Application US/09640198D
; Patent No. 6586411
; GENERAL INFORMATION:
; APPLICANT: Russell, Stephen
; APPLICANT: Kay Whye, Peng
; TITLE OF INVENTION: System for Monitoring the Location of
; FILE REFERENCE: 07039-295001
; CURRENT APPLICATION NUMBER: US/09/640,198D
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 60/149,168
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-640-198D-4

Query Match 3.7%; Score 108; DB 4; Length 618;
Best Local Similarity 19.1%; Pred. No. 0.042;
Matches 103; Conservative 71; Mismatches 160; Indels 204; Gaps 27;
Qy 14 CGSLADYLTSAKFLYLHLSLSTWGDMMHFAVSFLVELYGNLSLLTAVYGLVVGASVL 73
Db 132 CGTL-QYLVAI--MLYTG-----IVIYAPALILNQVTGLDIWASLL 169
Qy 74 VLGAIGDWDKNAKLVKVAQTSLVQNVSVILCGIILMMVFLHKLHLLTMVHGWV-LTSC 132
Db 170 STGLICTLYTVGKMAVWTD--VFQVVMVVG-----WVILARG 209
Qy 133 YILITITIANIANLASTATAITQ-----RDWIVVVG-----164
Db 210 VILGGPRNVLSLAQNSHRLNMBDDPDRSRYTFWTFIVGTLVLSWYGVNQVQRY 269
Qy 165 -----EDRSKLANMNATIRRIDLTNLAIPMAVGQIMTFGSPVTCGFSIGNLVSMCVE 219
Db 270 VACHTEGAKLAL-----VNQL-----GLFLIVASAA-CG-----300
Qy 220 YVLLMKVYOK--TPALAVKAGLKEEETELKOLNHLKDTKPK-----LEGTH 264
Db 301 --IVMFYKCDPILLTGRISAPQYMPDLLVLDIFEDLPGVGLFLACAYSGTILSTASTS 358
Qy 265 LMVGRKDSNIHELEHQEPTCASQMAEPRTFRDGNVSYNPOVFVLAGMGLAFLYMTVLGF 324
Db 359 INAMAATVEDLIKPRMPGLAPR-----KLVFIS-KGLSFIY-----GS 396
Qy 325 DCITTYAYTQGLSGSILMGLASAITG---IMGTV-----AFT--WLRRKQGL--VR 371
Db 397 ACLTV-----AALSSLLGGVLOGSFTVMGVISGPLLGAFTLGMLLPACNTPGVL 446

Qy 372 TGLISGLAQLSCLILCVISVMPGSPDLDSVPEDIRSRFIQGESITPTKIPBITTEIY 431
Db 447 SGLAAGLA-VSLWAVGATLYPPGE-----QTMGVLP TSAAGCTNDSV 488
Qy 432 M---SNGSNANIVPETSPE-----SVPIISVLLFAGVIAARIGLWSPDLTWTQLL 480
Db 489 LLGPPGATNASGIPSSGMDTGRPALADTFVAISLYLYGALG-----TLTMTL 536
RESULT 11
US-09-134-001C-3299
; Sequence 3299, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3299
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3299

Query Match 3.7%; Score 107; DB 4; Length 411;
Best Local Similarity 19.7%; Pred. No. 0.028;
Matches 96; Conservative 70; Mismatches 167; Indels 154; Gaps 20;
Qy 17 LADYLTSAKFLYLHLSLSTWGDMMHFAVSFLVELYGNLS-----LLLTAV 63
Db 19 MAKYFPSSFLFLGNWIGVG-----LWVFLTYHNNAVYGLVNFCLIFILLLSV 71
Qy 64 YGLVAGSVLVGAIIGDWDKNAKLVKVAQTSLVQNVSVILCGIILMMVFLHKLHLLTM 123
Db 72 W-----AGS-----IADKYDKGLLRITITSSFI--ITALLC-----LLT- 104
Qy 124 YHGVWLTSCYILII---TIANIANLASTATAITQSDWIVVWVAGEDRSKLANMNATIRR 179
Db 105 ---YLSNLSIFILLYATFRGILSAVETPVROAVLPDL-----SSKISTTQAVSFH 153
Qy 180 ---IDQLTNILAPMAVGQIMTFGSPV-----ICGFTSGWNLVSMCVYVLLMKVYQKTP 231
Db 154 SFIINICRSIGPAIAGSLIAVYHTPTTFLAQAVCYFIAVLCIPHFVILSKEGRALP 213
Qy 232 ALAVKAGLKEEETELKOLNHLKDTKPKLEGTHLMGVKOSNIHELEHQEPTCASQMAEP 291
Db 214 LKV-----LNYFKSN-----LEGSQ-----229
Qy 292 FRFRDGNVSYNQVPFLAG---MGLAFLYMTVLGPDCTITTYAYTQGLSGSILSILMGA 348
Db 230 -----IFITSIIIMATGFSYTTVLP---VLTNHFPG--OSQVFGIATMF 269
Qy 349 SAITGIMGTVAFTWLARKKCGLVR---TGLISGLAQLSCLILCVISVMPGSPDLDSVSP 404
Db 270 CAIGGIVATIVLSILKHLSTVKMYLSSILFGIALLGIIHHLVVMFICITLIGL-FSQ 328
Qy 405 FEDIRSRFIQGESITPTKIPBITTEIYMSNG-----SNSANIVPETSPEVPI 452
Db 329 WARTTNRVYFQHSVKDCDRGKVLIIIMDRGMIPGLSLIMSFFADMEGILTTFTMGIST 388
Qy 453 ISVSLLF 459
Db 389 ISISLIF 395

RESULT 12

US-09-328-352-5222
; Sequence 5222, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5222
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5222

Query Match 3.7%; Score 107; DB 4; Length 465;
Best Local Similarity 20.0%; Pred. No. 0.034;
Matches 85; Conservative 74; Mismatches 140; Indels 126; Gaps 21;
QY 59 LLTAVYGLVAGSVLVIG-AIIGVDVKNARKVAQTSVLVQNVSVILGIIIMVFLHK 117
DB 131 VMAAIFGVWAGSFIEMGVSGILPWVK-----LITPLVTGIVVLLIGLTLI---K 178
QY 118 HELTMYHGWLTSCYILIIITIANIANLASTATAITQRDWIVVAGEDRSKLANMNATI 177
DB 179 EGLISMGGOAMQDH-----TFASADNLMSCTVLAI-----IIVL---NRIRVWIKSSA 227
QY 178 RRIDQLTNILAPMAVGQIMT-----FGSPVIGC-----GFISGNLVSVCVEY 220
DB 228 -----ILIALVIGYLAGFMGYLDFSGLDKAPVLIQITPMHFGLSFSGWL-----F 273
QY 221 VLLWKVYQKTPALAVKAGLKEEETELKQLNLHKDTEPKLEGTHLM-----GVKDSNIHE 275
DB 274 IPNAFIVLVSLEAI-----GDTVATSKLS-----NOPVDGQWQRIKGGVLVNGANS 322
QY 276 LEHEQETCASQAEPRTRFDGHWVSYNQPVFLAGGLAFL-----YMTVLGFCIT 329
DB 323 L-----LAGLNTFPSSVFAQNGVIGLTGVASRYGVGWIATALLILGLFPFAVA 371
QY 330 G--YAYTQGLSGSILSILMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLILC 387
DB 372 GVTAQVPAVLGGAVMFGVAASGI-----NILSSI-HLDRALL 412
QY 388 VISVFMGSPDLDSVSPFDIRSRFQGESITPKIPEITEIYMS-----NGNSANIVPE 444
DB 413 IIAISL---ALGLGAQVPQILEH-----LPLEFKNIFSSGVATGGIAALILNV 458
QY 445 TSPES 449
DB 459 VLPET 463

RESULT 13
US-09-031-392-4
; Sequence 4, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D.; Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-4
Query Match 3.6%; Score 104; DB 2; Length 534;
Best Local Similarity 20.1%; Pred. No. 0.086;
Matches 89; Conservative 73; Mismatches 126; Indels 154; Gaps 24;
QY 42 WIFAVSVFLVELYGNLSLLTAYVGLVAGSVLVGLAIGDWDV-DKNARKVAQTSVLVQNV 100
DB 109 WLSVSMF-----AVGMVSSFTV-----GWIGDLGRVK-----AMLVNV 144
QY 101 VSVILCGIILMVLHKLHLLTMYHGWLTSCYILIIITIANIANLASTATAITQRDWIV 160
DB 145 VLSIAGNLWLGAKMPSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
QY 161 VVAGEDRSKLANMNATIRIDQLTNLAPMAVGQIMTFGSPVIGCGFISG-----WNL--- 213
DB 187 YVSEVSPALRGALGTLHQLAIVTGLI-----SQVLGLDFLLGNDELWPLLLG 235
QY 214 ---VSMCVYVLLWKVYQKTPALAVKAG-LKEEETELKQLNLHKDTEPKLEGTHLMGVK 269
DB 236 LSGVAALLQFFLLLCPESPRYLYIKLGKVEBAKSLKELGNCD---PMK----- 283
QY 270 DSNIEHEHEOBTCTAS-----QWAEPPFRTRFDGHWVSYNQ 305
DB 284 --EIAEMEKEKEAEASEKRVSGQLFSSKYQAVIVALMVQISOQFSGINA--IFYYST 339
QY 306 PVF-LAGMGLAFLYMTVLGFCITTYAYTQGLSGSILSILMGASA-----ITGIMG-- 356
DB 340 NIFQAGVGPVYVYATI-GVGVVNTVF-----TVISVFLVEKAGRRSLFLAGLMGL 390
QY 357 --TVAFWTWLRKCGLVRTGLISGLAQLSCLILCVISVFM-----PG----- 395
DB 391 ISAVAMT-----VGLV---LLSQFAMWSYVMVAIFLVFIFFEVGPGPIFWFIVAEFLSQ 442
QY 396 --SPDLDSVSPEDIRSRFIOG 415
DB 443 GPRPAIAVAGFCNACNFIVG 464

RESULT 14
US-09-299-549-4
; Sequence 4, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston

STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-APR-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Meikiejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-4

Query Match 3.6%; Score 104; DB 3; Length 534;
Best Local Similarity 20.1%; Pred. No. 0.086;
Matches 89; Conservative 73; Mismatches 126; Indels 154; Gaps 24;
Qy 42 WHPAVSVFLVLYGNSLLLTAVYGLVAGSVLVGAIIGDWV-DKNARKLVAQOTSLVVQN 100
Db 109 WLSVSMF-----AVGVMVSSFTV-----GWIGDLGRVK-----AMLVN 144
Qy 101 VSVILCGIILMMVFLHKKHELLTMYHGWLTSCYILITIANIANLASTATAITQDWIV 160
Db 145 VLSIAGNLLMGLAKMGPSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
Qy 161 VVAGEDSKLANMNAIRRIDQLTNLA PMAVGQIMTFGSPVIGCGFISG-----WNL--- 213
Db 187 YVSEVSPALRGALGTLHQAIVTGILI-----SQVLGLDFLLGNDELWPLLLG 235
Qy 214 ---VSMCEVYVLMKVYKTPALAVKAG-LKEBETELKQLNLHKTPEKPLEGTHLMGVK 269
Db 236 LSGVAALLQFLLCPESPRLYIKLGKVEAKSKLRGNC-----PMK----- 283
Qy 270 DSNIEHEHEQPTCAS-----QMAEPRTFRDGVSVYNNQ 305
Db 284 --EIAEMEKEQEAASEKRVISIGLFSKKYRQAVIVALMVQISQQFSGINA--IFYST 339
Qy 306 PVF-LAGMGLAFLYMTVLGPDCTTGYAYTQGLSGSILSLMGASA-----ITGIMG-- 356
Db 340 NIFQAGVGQPVYIATI-GVGVNVTVF-----TVISVFLVEKAGRRSLFLAGLMGL 390
Qy 357 --TVAFWLRKCGLVRTGLISGLAQLSCLILCVISFVM-----PG----- 395
Db 391 ISAVAMT-----VGLV-----LLSQFAWMSYVSMVAIFLVIFFEVFGPGPIPWFI VAE LFSQ 442
Qy 396 --SPLDLSVSPFDIRSRFIQ 415
Db 443 GPRPAATAVAGFCNACNFIVG 464

RESULT 15

US-09-610-417-4
Sequence 4, Application US/09610417
Patent No. 6346374
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.

Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/610,417
FILING DATE: 05-Jul-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/299,549
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meikiejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-610-417-4

Query Match 3.6%; Score 104; DB 4; Length 534;
Best Local Similarity 20.1%; Pred. No. 0.086;
Matches 89; Conservative 73; Mismatches 126; Indels 154; Gaps 24;
Qy 42 WHPAVSVFLVLYGNSLLLTAVYGLVAGSVLVGAIIGDWV-DKNARKLVAQOTSLVVQN 100
Db 109 WLSVSMF-----AVGVMVSSFTV-----GWIGDLGRVK-----AMLVN 144
Qy 101 VSVILCGIILMMVFLHKKHELLTMYHGWLTSCYILITIANIANLASTATAITQDWIV 160
Db 145 VLSIAGNLLMGLAKMGPSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
Qy 161 VVAGEDSKLANMNAIRRIDQLTNLA PMAVGQIMTFGSPVIGCGFISG-----WNL--- 213
Db 187 YVSEVSPALRGALGTLHQAIVTGILI-----SQVLGLDFLLGNDELWPLLLG 235
Qy 214 ---VSMCEVYVLMKVYKTPALAVKAG-LKEBETELKQLNLHKTPEKPLEGTHLMGVK 269
Db 236 LSGVAALLQFLLCPESPRLYIKLGKVEAKSKLRGNC-----PMK----- 283
Qy 270 DSNIEHEHEQPTCAS-----QMAEPRTFRDGVSVYNNQ 305
Db 284 --EIAEMEKEQEAASEKRVISIGLFSKKYRQAVIVALMVQISQQFSGINA--IFYST 339
Qy 306 PVF-LAGMGLAFLYMTVLGPDCTTGYAYTQGLSGSILSLMGASA-----ITGIMG-- 356
Db 340 NIFQAGVGQPVYIATI-GVGVNVTVF-----TVISVFLVEKAGRRSLFLAGLMGL 390
Qy 357 --TVAFWLRKCGLVRTGLISGLAQLSCLILCVISFVM-----PG----- 395
Db 391 ISAVAMT-----VGLV-----LLSQFAWMSYVSMVAIFLVIFFEVFGPGPIPWFI VAE LFSQ 442
Qy 396 --SPLDLSVSPFDIRSRFIQ 415

Db 443 GPRPAIAVAGFCNWACNFIVG 464

Search completed: November 5, 2003, 18:37:14
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:32:22 ; Search time 45 Seconds
(without alignments)
2179.308 Million cell updates/sec

Title: US-09-715-927-6
Perfect score: 2929
Sequence: 1 MTRAGDHNRRGCCGLADY.....ACGPDAKEVRKENQANTSW 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:**
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2924	99.8	571	11	US-09-973-180-2
2	2924	99.8	571	11	US-09-973-180-3
3	2922	99.8	571	11	US-09-973-180-4
4	2917	99.6	571	9	US-09-729-674-130
5	1789	61.1	383	11	US-09-809-391-460
6	1789	61.1	383	12	US-09-882-171-460
7	125.5	4.3	575	15	US-10-156-761-10350
8	120	4.1	458	12	US-10-283-423-6
9	120	4.1	458	12	US-10-213-821-6
10	119	4.1	537	10	US-09-738-626-6730
11	114.5	3.9	424	10	US-09-738-626-3641
12	114	3.9	392	12	US-09-769-744A-120
13	113	3.9	522	9	US-09-804-551B-38
14	108.5	3.7	486	12	US-10-419-190-2
15	108	3.7	618	10	US-09-995-007-2

16	107.5	3.7	488	9	US-09-815-242-10767	Sequence 10767, A
17	104	3.6	534	10	US-09-981-947A-4	Sequence 4, Appli
18	104	3.6	608	15	US-10-156-761-14861	Sequence 14861, A
19	103.5	3.5	471	9	US-09-989-861-17	Sequence 17, Appl
20	102.5	3.5	473	10	US-09-746-491-72	Sequence 72, Appl
21	102.5	3.5	529	14	US-10-051-902-28	Sequence 28, Appl
22	102.5	3.5	529	14	US-10-051-909-28	Sequence 28, Appl
23	101.5	3.5	426	15	US-10-187-267A-43	Sequence 43, Appl
24	101.5	3.5	435	15	US-10-156-761-11512	Sequence 11512, A
25	101.5	3.5	739	12	US-10-328-194A-3	Sequence 3, Appli
26	101.5	3.5	739	15	US-10-205-823-369	Sequence 369, App
27	101	3.4	455	9	US-09-815-242-13794	Sequence 13794, A
28	101	3.4	551	14	US-10-216-355-4	Sequence 4, Appli
29	100.5	3.4	411	9	US-09-828-746-6	Sequence 6, Appli
30	100.5	3.4	412	15	US-10-156-761-9945	Sequence 9945, Ap
31	100.5	3.4	506	9	US-09-864-761-46512	Sequence 46512, A
32	100.5	3.4	1356	15	US-10-090-280-33	Sequence 33, Appl
33	100.5	3.4	1359	15	US-10-090-280-34	Sequence 34, Appl
34	98.5	3.4	411	9	US-09-828-746-2	Sequence 2, Appli
35	98.5	3.4	411	15	US-10-121-746-83	Sequence 83, Appl
36	98.5	3.4	500	12	US-10-247-813-6	Sequence 6, Appli
37	98.5	3.4	500	12	US-10-247-813-28	Sequence 28, Appl
38	98	3.3	458	15	US-10-001-073-46	Sequence 46, Appl
39	98	3.3	459	10	US-09-738-626-6998	Sequence 6998, Ap
40	98	3.3	461	15	US-10-225-567A-44	Sequence 44, Appl
41	98	3.3	462	15	US-10-001-073-44	Sequence 44, Appl
42	97.5	3.3	338	10	US-09-886-055-205	Sequence 205, App
43	97.5	3.3	338	11	US-09-804-291-205	Sequence 205, App
44	97.5	3.3	338	12	US-10-017-161-30	Sequence 30, Appl
45	97.5	3.3	808	15	US-10-223-047-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-973-180-2
; Sequence 2, Application US/09973180
; Publication No. US20030082553A1
; GENERAL INFORMATION:
; APPLICANT: Oestra, Ben Heutink, P Duijin, C.M.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR AUTOSOMAL DOMINANT HEMOCHROMATOSIS
; FILE REFERENCE: 24584
; CURRENT APPLICATION NUMBER: US/09/973,180
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/301,429
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-180-2

Query Match	99.8%	Score	2924	DB	11	Length	571
Best Local Similarity	99.8%	Pred. No.	3.7e-272				
Matches	570	Conservative	1	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MTRAGDHNRRGCCGLADYLTSAKFLLYLGHSLSTWGDNRMHFAVSFLVELYGNLSLL	60				
Db	1	MTRAGDHNRRGCCGLADYLTSAKFLLYLGHSLSTWGDNRMHFAVSFLVELYGNLSLL	60				
Qy	61	TATVGLVAGSVLVILGAIIGDWDKNARLKVQATSLVVQNVSVILCGIILMMVFLKHKL	120				
Db	61	TATVGLVAGSVLVILGAIIGDWDKNARLKVQATSLVVQNVSVILCGIILMMVFLKHKL	120				
Qy	121	LTWVHGWLVSCTLIITITIANIANLASTATAITQTDWIVVWVAGEDSKLANMATTIRI	180				
Db	121	LTWVHGWLVSCTLIITITIANIANLASTATAITQTDWIVVWVAGEDSKLANMATTIRI	180				
Qy	181	DQLTNILAPVAGQIMTFSPGICGFIGSNLVSMCVYLLMKVYQKTPALAVKAGLK	240				
Db	181	DQLTNILAPVAGQIMTFSPGICGFIGSNLVSMCVYLLMKVYQKTPALAVKAGLK	240				

Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
Qy 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTVA 360
Qy 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMFGSPDLDSVSPFEDIRSRFIQGESITP 420
Db 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMFGSPDLDSVSPFEDIRSRFIQGESITP 420
Qy 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Qy 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

RESULT 2
US-09-715-927-6-3
; Sequence 3, Application US/09973180
; Publication No. US20030082553A1
; GENERAL INFORMATION:
; APPLICANT: Oestra, Ben Heutink, P Duijin, C.M.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR AUTOSOMAL DOMINANT HEMOCHROMATOSIS
; FILE REFERENCE: 24584
; CURRENT APPLICATION NUMBER: US/09/973,180
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/301,429
; PRIOR FILING DATE: 2001-06-29
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 144
; OTHER INFORMATION:
US-09-715-927-6-3

Query Match 99.8%; Score 2924; DB 11; Length 571;
Best Local Similarity 99.8%; Pred. No. 3.7e-272;
Matches 570; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Db 1 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Qy 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLVQATSLVQVQNSVILCGIILMVFLKHHEL 120
Db 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLVQATSLVQVQNSVILCGIILMVFLKHHEL 120
Qy 121 LTMVHGWLVTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Db 121 LTMVHGWLVTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Qy 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
Qy 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300

Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTVA 360
Qy 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMFGSPDLDSVSPFEDIRSRFIQGESITP 420
Db 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMFGSPDLDSVSPFEDIRSRFIQGESITP 420
Qy 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Qy 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

RESULT 3
US-09-715-927-6-4
; Sequence 4, Application US/09973180
; Publication No. US20030082553A1
; GENERAL INFORMATION:
; APPLICANT: Oestra, Ben Heutink, P Duijin, C.M.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR AUTOSOMAL DOMINANT HEMOCHROMATOSIS
; FILE REFERENCE: 24584
; CURRENT APPLICATION NUMBER: US/09/973,180
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/301,429
; PRIOR FILING DATE: 2001-06-29
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 144
; OTHER INFORMATION: Xaa = any amino acid except Asp
US-09-715-927-6-4

Query Match 99.8%; Score 2922; DB 11; Length 571;
Best Local Similarity 99.8%; Pred. No. 5.8e-272;
Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Db 1 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Qy 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLVQATSLVQVQNSVILCGIILMVFLKHHEL 120
Db 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLVQATSLVQVQNSVILCGIILMVFLKHHEL 120
Qy 121 LTMVHGWLVTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Db 121 LTMVHGWLVTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Qy 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
Qy 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTVA 360

Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVF 360
 Qy 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSGLDLSVSPEDIRSRFIQESITP 420
 Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSGLDLSVSPEDIRSRFIQESITP 420
 Qy 421 TKIPITTEIYMSNGSANSIVPETSPEPVIIISVLLFAGVIAAIGLWSFDLTVTQLL 480
 Db 421 TKIPITTEIYMSNGSANSIVPETSPEPVIIISVLLFAGVIAAIGLWSFDLTVTQLL 480
 Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
 Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
 Qy 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571
 Db 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571

RESULT 4
 US-09-729-674-130
 ; Sequence 130, Application US/09729674
 ; Patent No. US20010039335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: LaVallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merbers, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steindinger II, Robert J.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fectel, Kim
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6055-64X
 ; CURRENT APPLICATION NUMBER: US/09/729,674
 ; CURRENT FILING DATE: 2000-12-04
 ; PRIOR APPLICATION NUMBER: 09/539,330
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 283
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 130
 ; LENGTH: 571
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (202)
 ; NAME/KEY: UNSURE
 ; LOCATION: (504)
 ; US-09-729-674-130

Query Match 99.6%; Score 2917; DB 9; Length 571;
 Best Local Similarity 99.6%; Pred. No. 1.7e-271;
 Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MTRAGDHNQRQCCGSLADYLTSAKFLYLGLHSLSTWGDNRWHFAVSFLVELYGNLSLL 60
 Db 1 MTRAGDHNQRQCCGSLADYLTSAKFLYLGLHSLSTWGDNRWHFAVSFLVELYGNLSLL 60
 Qy 61 TAVYGLVAVGSVLGAILGVDKRNARLKAQTSLVVQNSVILICGILMMVFLKHREL 120
 Db 61 TAVYGLVAVGSVLGAILGVDKRNARLKAQTSLVVQNSVILICGILMMVFLKHREL 120
 Qy 121 LTMHGWLTSCYILIIITIANIANLASTATAITQRDWIVVAVGDRSKLANMATIRRI 180
 Db 121 LTMHGWLTSCYILIIITIANIANLASTATAITQRDWIVVAVGDRSKLANMATIRRI 180

Qy 181 DQNLNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
 Db 181 DQNLNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
 Qy 241 BEETELQNLNKHDTPEKPLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWV 300
 Db 241 BEETELQNLNKHDTPEKPLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWV 300
 Qy 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVF 360
 Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVF 360
 Qy 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSGLDLSVSPEDIRSRFIQESITP 420
 Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSGLDLSVSPEDIRSRFIQESITP 420
 Qy 421 TKIPITTEIYMSNGSANSIVPETSPEPVIIISVLLFAGVIAAIGLWSFDLTVTQLL 480
 Db 421 TKIPITTEIYMSNGSANSIVPETSPEPVIIISVLLFAGVIAAIGLWSFDLTVTQLL 480
 Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
 Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
 Qy 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571
 Db 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571

RESULT 5
 US-09-809-391-460
 ; Sequence 460, Application US/09809391
 ; Publication No. US20030049618A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: 186 Human Secreted proteins
 ; FILE REFERENCE: P2002P2
 ; CURRENT APPLICATION NUMBER: US/09/809,391
 ; CURRENT FILING DATE: 2001-03-16
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 761
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 460
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (383)
 ; OTHER INFORMATION: Xaa equals stop translation
 ; US-09-809-391-460

Query Match 61.1%; Score 1789; DB 11; Length 383;
 Best Local Similarity 92.4%; Pred. No. 2.8e-163;
 Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;
 Qy 190 MAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEEETELKQL 249
 Db 1 MAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEEETELKQL 60
 Qy 250 NLHKDTEPKLEGTLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWVSYNQPVFL 309
 Db 61 NLHKDTEPKLEGTLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWVSYNQPVFL 120
 Qy 310 AGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVFTWLRKCGL 369
 Db 121 AGMGLAFYMTVLGFDCTTGYAYTQGLSGFHPQYFDGSIYNNNGNSCYLATSQWVF 180
 Qy 370 VRTGLISGLAQSLCLILCVISVFMPSGLDLSVSPEDIRSRFIQESITPTKIPETITE 429
 Db 181 GSAGLISGLAQSLCLILCVISVFMPSGLDLSVSPEDIRSRFIQESITPTKIPETITE 240
 Qy 430 IYMSNGSANSIVPETSPEPVIIISVLLFAGVIAAIGLWSFDLTVTQLLQENVIESER 489

Db 241 IYNSGNSANIWPETSPESVPIISVLSLLFAGVIARIGLWSTDLTQTLLQENVIESR 300
QY 490 GIINGVQNSMYLLDLLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFRFAQNTLGNK 549
Db 301 GIINGVQNSMYLLDLLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFRFAQNTLGNK 360
QY 550 LFACGPDPAKEVRKENQANTSVV 571
Db 361 LFACGPDPAKEVRKENQANTSVV 382
RESULT 6
US-09-882-171-460
; Sequence 460, Application US/09882171
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22

Query Match 61.1%; Score 1789; DB 12; Length 383;
Best Local Similarity 92.4%; Pred. No. 2.8e-163;
Matches 353; Conservative 3; Mismatches 26; Indels 0

RESULT 7

Query Match 4.3%; Score 125.5; DB 15; Length 575;
Best Local Similarity 20.4%; Pred. No. 0.0049;
Matches 102; Conservative 70; Mismatches 162; Indels 165; Gaps 23;

39	QY	DRMWFHFAVSFVLVELYGNSSLTATA-----VYGLVAVGSVLVLGAAIGDWMVDKNA	87
		: : : :	
57	Db	DLQW--VIDGYTPVPAALMLLIGALGDKYSRRGALVGLGVFEG--CAVAGSLVDSAT	110
		: : : :	
88	QY	RLKVAQTSLVQVNSVILCGIILMMVFLKH-----LLTM	123
		: : : :	
111	Db	GVIARAVNGVGAALIMPATLSLAATPPRAERAKAITLWTATAGLATAAGSLVAGALLQ	170
		: : : :	
124	QY	YHGVLVLTSCVILLITITANIANLASTATAITIQDMLVVVAGEDRSKLANMNATIRIDOL	183
		: : : :	
171	Db	NHGW--SSTFLNPVIAFVAVIGT-----FVLVPPSRA-----AHHRIDIV	210
		: : : :	
184	QY	THILAPMAVGQI--MTFGSPVIGCGFISGWNVLVSMCVVEYVLL--WKVYQKTPALAVKAGL	239

Db 211 GGLSVVWIGSLVYMIIEGPHFGVGVKAVTAAVAAGVGVFLFIGWELRHRPVL----- 264
Qy 240 KEBETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHQETPCASQMAPEPRTFRDGW 299
Db 265 -----DVRFAHRR--FAGSNLA----- 280
Qy 300 VSYNQPVFLAGGLAFLYMT-----VLGFDCITTTGAYVYQGLSGSTLSILMGASAITGI 354
Db 281 VALF-----FLAVFG-ABEYLTQHQLQFLVGYDALDTGVRMLP-LAGAVFV-----GSALTG- 329
Qy 355 MGTVAFTWLRKCGCL-----VRLTGISGLAQLSCILC-----VISVFMPSGLDLSV 402
Db 330 -----YLTFRVGMKITVAGVMVGGTAALALLTRVDAASSYGGDFVLPLVILGLAIGLAL 382
Qy 403 SPED-IRSPICGESITPKIPEITTEIYMS-----NGSNSANIVPTSPESVP 451
Db 383 SPCTDAIMGAPPEALGVGAVNDTSLGGLGIALGSLVLAGSYSSHLSDATTSKLP 442
Qy 452 IISVS-----LLFAGVIAAR 466
Db 443 ASALSQAQDSVGGAGYAVAR 461

RESULT 8

US-10-283-423-6
; Sequence 6, Application US/10283423.
; Publication No. US20030162223A1
; GENERAL INFORMATION:
; APPLICANT: Lowery, David E.
; APPLICANT: Smith, Valdin G.
; APPLICANT: Kubiak, Teresa M.
; APPLICANT: Larsen, Martha J.
; TITLE OF INVENTION: Drosophila G Protein Coupled Receptors, Nucleic Acids, And Method
; TITLE OF INVENTION: Related To The Same
; FILE REFERENCE: PHRM0002-102
; Application Project
; -----
; CURRENT APPLICATION NUMBER: US/10/283,423
; CURRENT FILING DATE: 2002-10-30
; Earlier Applications
; -----
; PRIOR APPLICATION NUMBER: PriorAppNumber : 09/693,746
; PRIOR FILING DATE: PriorFilingDate : 2000-10-20
; PRIOR APPLICATION NUMBER: PriorAppNumber : 09/425,676
; PRIOR FILING DATE: PriorFilingDate : 1999-10-22
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 458
; TYPE: PRT
; ORGANISM: D. melanogaster
US-10-283-423-6

Query Match 4.1%; Score 120; DB 12; Length 458;
Best Local Similarity 20.1%; Pred. No. 0.012;
Matches 77; Conservative 60; Mismatches 98; Indels 148; Gaps 21;

Qy 5 GDHNRQCGCGSLADYLSAKFLLYLG-----HSLSTWG----- 38
Db 12 GEH-----LSGYASSNSVRYLDDRHPLDYLDTGTVHALNTTAINSTDNETGSRP 62
Qy 39 -----DRMWHFAVSFVFLVELYGNLSLLTAVYGLVAVGASVVLVGAIGDWW 83
Db 63 LDPVLIDRFLSNRVDSFWYH-----MLISMYG-----VLIVFGAL--GNTLWIAVIRKPI 112
Qy 84 DKNARLKVAQTSLVVQNVSV---ILCGIILMMVFLHKHELLTMYHGWLTSCYILITIA 140
Db 113 MPTAR-----NLFILNLAISDLLCLVTPMLTLM---EILSKY--WPYGSCSILCKTIA 161
Qy 141 NIANL-----ASTATAITIQDMIVVVGEDRSKLANMNATIRRIDQLTNIIA-PMAV-G 193
Db 162 MLQALCIVFSTISITAIAFDQYVIVYPTDRSLQFVGAVTILAGIWWALLASPLFVYK 221

Qy 194 QIMTGGSPV---IG-----CGFISGW-----NLVSMCIVEY-----VLLWK 225
Db 222 EINTDTPALLOQIGLQDTIPLYC--IEDWPSRNGRFYISFSLCQVLYLPILIVSVAYFG 279
Qy 226 VYQKTPA-----LAVKAGLKEEETE-----LKQNLHKDTE 256
Db 280 IYNKLKSRITTVAVQASSAQRKVERGRMRKRTNCLLSIAITIFGVSWLPLNFFNLVADME 339
Qy 257 PKPLEGT-----HLMGVKDS 271
Db 340 RSPVTQSMVLRYAICHMIGMSSA 362

RESULT 9

US-10-213-821-6
; Sequence 6, Application US/10213821
; Publication No. US20030180297A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; APPLICANT: Lowery, David E.
; APPLICANT: Smith, Valdin G.
; APPLICANT: Kubiak, Teresa M.
; APPLICANT: Larsen, Martha J.
; TITLE OF INVENTION: Drosophila G Protein Coupled Receptors, Nucleic Acids, And Methods
; TITLE OF INVENTION: Related To The Same
; FILE REFERENCE: 6297.2Cp
; CURRENT APPLICATION NUMBER: US/10/213,821
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 09/693,746
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/425,676
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 458
; TYPE: PRT
; ORGANISM: D. melanogaster
US-10-213-821-6

Query Match 4.1%; Score 120; DB 12; Length 458;
Best Local Similarity 20.1%; Pred. No. 0.012;
Matches 77; Conservative 60; Mismatches 98; Indels 148; Gaps 21;

Qy 5 GDHNRQCGCGSLADYLSAKFLLYLG-----HSLSTWG----- 38
Db 12 GEH-----LSGYASSNSVRYLDDRHPLDYLDTGTVHALNTTAINSTDNETGSRP 62
Qy 39 -----DRMWHFAVSFVFLVELYGNLSLLTAVYGLVAVGASVVLVGAIGDWW 83
Db 63 LDPVLIDRFLSNRVDSFWYH-----MLISMYG-----VLIVFGAL--GNTLWIAVIRKPI 112
Qy 84 DKNARLKVAQTSLVVQNVSV---ILCGIILMMVFLHKHELLTMYHGWLTSCYILITIA 140
Db 113 MPTAR-----NLFILNLAISDLLCLVTPMLTLM---EILSKY--WPYGSCSILCKTIA 161
Qy 141 NIANL-----ASTATAITIQDMIVVVGEDRSKLANMNATIRRIDQLTNIIA-PMAV-G 193
Db 162 MLQALCIVFSTISITAIAFDQYVIVYPTDRSLQFVGAVTILAGIWWALLASPLFVYK 221
Qy 194 QIMTGGSPV---IG-----CGFISGW-----NLVSMCIVEY-----VLLWK 225
Db 222 EINTDTPALLOQIGLQDTIPLYC--IEDWPSRNGRFYISFSLCQVLYLPILIVSVAYFG 279
Qy 226 VYQKTPA-----LAVKAGLKEEETE-----LKQNLHKDTE 256
Db 280 IYNKLKSRITTVAVQASSAQRKVERGRMRKRTNCLLSIAITIFGVSWLPLNFFNLVADME 339
Qy 257 PKPLEGT-----HLMGVKDS 271
Db 340 RSPVTQSMVLRYAICHMIGMSSA 362

RESULT 10

US-09-738-626-6730
; Sequence 6730, Application US/09738626
; Publication No. US20020197605A1

GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6730
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6730

Query Match 4.1%; Score 119; DB 10; Length 537;

Best Local Similarity 20.9%; Pred. No. 0.019;
Matches 122; Conservative 72; Mismatches 172; Indels 218; Gaps 29;

Qy	35	STWGRMWH-----FAVSFLVELYGNLSLLTAVYGLVAGSV-----LVLG 76
Db	5	STWSRTWHRKASRPVSILWV-----LVIAIHLPLPEYRWLHLFTLG 51
Qy	77	AIIGD-----WVDKNARLVAQTSLVVQ--NVSVILCGIILMMVFLHKLHLLTYHGW 127
Db	52	AITNSIVVMSQHFTEKFLHLKLESKRPAQLKIRLVNGLIIVTIG-----QMIGQW 104
Qy	128	VLTSYILII-----TIANIANLAS-----147
Db	105	IVTSVGATIVGGALAWHAGSLASQFSRAKQGPFAAVIAYVASACCLPFGAFAGALLSK 164
Qy	148	-----TATAITIQRDWIVVAVAGEDRSKLANMNATIRIDQLTNILAPWVQOI 195
Db	165	ELSHGLQBRVLLTHVI-----NFGVGFAGALGSLVLFAAIWR--TKIRHNFTPSVVG-I 218
Qy	196	MTFGSPVIGCGFI--SGNVLVSMCVYVLLWKVYQKTPALAVKAGLKEEETELKQLNLHK 253
Db	219	MAVSLPIIVTIGLLNNGVVAATGLAAVYAM-----LLAVMGWGRKASINLSFSTSTS 271
Qy	254	DTEPKPLEGTHLMGVKDSNIH--ELEHEQEBPTCASQMAEPRTFRDGMVSYNQPVFLAG 311
Db	272	TTAPLWLVTGLVLAQVAMHDGELYHVEVPTIA-----305
Qy	312	MGLAFYMTVLGFDCTTGYAYTOGLSGSILSLM--GASAI--TGIMGTVAFTWLRKCG 368
Db	306	-----LVIGF-----GAQLLIGWMSYLLPSTWGGGASAVRTG-----THLNTAG 345
Qy	369	LVRTGLISG-----LAQLSCLILCVISVFMFGSPGLDLSVFPFEDIRSRF 412
Db	346	LFRWTLLNGGLAIWLLTDSNWLRVVSVLSLIGALAVFVILLP-----KAVRAQ- 393
Qy	413	IQG-----ESITPKIP---EITTEIYM-----SNGSNSANIVP-ETSPESVPIISVS 456
Db	394	-RGVITKREPTTPPEPRLNQITAGISVLALILAAFGGLNPGVAPVASSNEDVYAVTIT 452

Qy

457 ----LLFAGVIAARIGLWSFDLTVTQLQENVIESERGIINGVQ 496

Db

453 AGDWMFIPDVIEWPAGK---SLEVTMLNEDDMVHDLK-FANGVQ 492

RESULT 11

US-09-738-626-3641
; Sequence 3641, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3641
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3641

Query Match 3.9%; Score 114.5; DB 10; Length 424;

Best Local Similarity 21.4%; Pred. No. 0.035;
Matches 86; Conservative 55; Mismatches 131; Indels 129; Gaps 16;

Qy	28	LYLGHSLSTWGRMWHFAVSFLVELYGNLS--LLLTAIVYGLVAGSVVLGAIIGDWYDK 85
Db	21	LWLGNAVATVIGAQLTVVAVPVQIYQMTGSSGYVGLTGLGLIPLVIFGLYGGSTADAFDK 80
Qy	86	NARLKVAQTSLVVQNVSVILCGIILMMVFLHKLHLLTYHGWVLT-----SCYILIIITI- 139
Db	81	RI-----VLICTTIGMCV-----TTAGFWVLITLGNENIMLLINFS 117
Qy	140	-----ANIANLASTATATIQRDWIVVAVAGEDRSKLANMNATIRIDQLTNILAPWVQIM 196
Db	118	LQOAFVAVNQPTTAIL--RSILPI-----DQLASATSLNMLLMQTAIVGPIIAGALI 169
Qy	197	TFGSPVIGCGFI--SGNVLVSMCVYVLLWKVYQ--KTPALAVKAGLKEEETELKQLNLH 252
Db	170	-----PLIGFWLFLDVSII--IPTLWAVNSLPKIPSGKVMKAG-----208
Qy	253	KDTEPKPLEGTHLMGVKDSNIHLEHEQEBPTCASQMAEPRTFRDGMVSYNQPVFLAGM 312
Db	209	-----FASVVDGLKYLQAGQVLL---226
Qy	313	GLAFYMTVLGFDCTTGYAYTOGLSGSILSLM--GASAI--TGIMGTVAFTWLRKCGLVRT 372
Db	227	-----MNVLDLIAMIFGMPRALYPEIAEVNFGGAGATM--LAFMYSSMAVGAVLG 277
Qy	373	GLISG-LAQLS-----CLILCVISVFMFGSPGLDLSVSP 404
Db	278	GVLSGWMVARISROGVAVYWCIIANGAAVALGG--VAIVVSP 316

RESULT 12

```

US-09-769-744A-120
; Sequence 120, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; APPLICANT: Hanebro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769, 744A
; PRIOR FILING DATE: 2001-01-26
; PRIOR FILING DATE: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 120
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-120

```

```

Query Match      3.9%; Score 114; DB 12; Length 392;
Best Local Similarity 20.1%; Pred. No. 0.035;
Matches 77; Conservative 67; Mismatches 132; Indels 108; Gaps 17;

QY 30 LGHSLSTGDRMWHFAVSVE---LVELYGNLSLLTAVYGLVAGSVLV--LGALIGDWVD 84
Db 17 MGNMLDYGNVULASMGTTGQVIGVQISLVT-----SLVNPFGVIS---D 64

QY 85 KNARLKVAQTSLVQNVSVILCGIILMMVFLHKLHLLTMVHGWLTSYLIITIANIAN 144
Db 65 RFSRKLMTADLV-----CGILCLAFIRND-----SNWIGA-----LIVANIVQ 106

QY 145 LASTATAITIQDWIVVAVAGEDRSKLANMNATIRIDQTNILAPMAVGQIMTFGSPVIG 204
Db 107 AIAPAFSRTANKAIIIEVEKED--EIVYINSRLSELVLQVVGVSPLVFLVLOFAS--- 160

QY 205 CGFTSGNVLVSMCVYVLLMKVQKTPALAVKAGLKEEETELKQLNLHDKTEPKPLEGTH 264
Db 161 -----LHMTLLDLSLTFPIAFVIVAFLPKEAKVQ-----EKKATGRD 199

QY 265 L-MGVKDSNIHELEHQE-----PTCASQMAEPFRTPRDGWSVYVQVFLAGMGLAFLYM 319
Db 200 IFVDIKD-GLHYIWHQOEIFFLLLVASSVNVFFFAAF-EFLLPFSNQL-----YGSEGAYA 252

QY 320 TVLGFDCITTYGAYTQGLSGSILMSILMGASAITGIMGTVAFTWLRKCGLVRTGLISGLA 379
Db 253 SIL-----TMGAIGSIIGALLASKIKANIVN----- 279

QY 380 QLSCLILCVISVFMFGSPDLDSVS 403
Db 280 -LILLALTGVGVFMGLPLPTFLS 302

```

```

RESULT 13
US-09-804-551B-38
; Sequence 38, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT APPLICATION NUMBER: US/09/804, 551B
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 38
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-804-551B-38

Query Match      3.9%; Score 113; DB 9; Length 522;
Best Local Similarity 20.0%; Pred. No. 0.067;
Matches 76; Conservative 58; Mismatches 98; Indels 148; Gaps 20;

QY 5 GDHNRQRCGCGSLADYLTSAKFLLYLG-----HSLSTWG----- 38
Db 59 GEH-----LSGVASSNSVRYLDRHPLDYLDTGVTVALNTTAINTSDLNETGSRP 109

QY 39 -----DRMWHFAVSFLVELYGNLSLLTAVYGLVAGSVLVGLAIGDWV 83
Db 110 LDPVLIDRELSNRAVDSPWYH---MLTSMYG---VLIVFGAL--GNTLVVIAVIRKPI 159

QY 84 DKNARLKVAQTSLVQNVSVILCGIILMMVFLHKLHLLTMVHGWLTSYLIITIANIA 143
Db 160 METAR-----NLFINLAISVTMPLTM-----EILSKY--WPYGSCSILCKTIAMQ 205

QY 144 NL-----ASTATAITIQDWIVVAVAGEDRSKLANMNATIRIDQTNILA-PMAV-GQIM 196
Db 206 ALCIFVSTISITAIADFQYQVIVPTDLSQFVGAVTILAGIWAALLLASPLFVYKELI 265

QY 197 TFGSPV---IG-----CGFISGW-----NLVSMCVVEY-----VLLWKVYQ 228
Db 266 NDTPTALLQOIGLQDTPYCY--IEDWPSRNGRFFYSIFSLCVQYLVLPILIVSVAYEGIN 323

QY 229 KTPA---LAVKAGLKEETE-----HLMGVKDS 271
Db 324 KUKSRITVAVVQSSAQKVERGRMKRTNCLLSIAIIFGVSWLPNFFNLVADMERSP 383

QY 260 LEGT-----HLMGVKDS 271
Db 384 VTQSMVRYVAICHMIGMSSA 403

```

```

RESULT 14
US-10-419-190-2
; Sequence 2, Application US/10419190
; Publication No. US20030166155A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: CL000872CON
; CURRENT APPLICATION NUMBER: US/10/419,190
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: 60/232,856
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/685,852
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Human
US-10-419-190-2

```

```

Query Match      3.7%; Score 108.5; DB 12; Length 486;
Best Local Similarity 22.3%; Pred. No. 0.16;
Matches 106; Conservative 65; Mismatches 176; Indels 129; Gaps 25;

QY 46 VSVFLVEL---YGNLSLLTA-VYGLVAGSVLVGLAIGDWVKNARKV-----AQ 93
Db 39 ISIFFVEFTQTYTQDYAQTAWIHSIV--DCVTMLCAPGSSVSNHLSQCVGMGLGLLAS 96

QY 94 TSLVYQNVSVILCGIILMMVFLHKLHLLTMVHGWLTS-----CYLIITIANIANLAST 148

```

Db 97 TGLILSSFATSL-----KHLXLT-----GYLTGLGFALCY--SPAIAWVKYFSR 140
Qy 149 ATAITIQDWIIVVAGEDRSKLANNNATIRIDQLTNIIAP---MAVGQIMTFGSPVIGC 205
Db 141 RKALA-----YGIAMSGSIG-----TFILAPVVQILLIQFSWRGALLILG 181
Qy 206 GFISGNLVMCEVYVLLWKVQKTPALAVKAGLKEETEELKQLNLHKDTPEKPLEGTHL 265
Db 182 GFV-----LNLGV-----CGALMRPITLKEDHTTPEQNHVCR-TQKEDIK----- 220
Qy 266 MGVKDSNIHELEHEOEPT---CASQMAEPRTFRDGVWYVYNQPVFLA-GWGLAFLYMTV 321
Db 221 ---RVSPSSLTKEWAQCLCCLQOEYSFLMSD-FVLAVSVLFMAYGCSPLFVIL-- 274
Qy 322 LGFDCITTYAVTQGLSGSILSILMGASAITGIMCTVAFTWLRK-----CGLVRT 372
Db 275 -----VPYALSVGVSHQQAFLMSILGVIDIIGNITFGMLTDRCLKNYQVVCYLEAV 327
Qy 373 GLISGLAQLSCLILCVISVFMFGSPDLDSVFPEDIRSRFIQGESITPTKIPETITEYM 432
Db 328 GM-DGLCYLCPLMQLSPLLPFS-----CTFGVFDGAYVTLIPVVTTEIVG 373
Qy 433 SNGNSA-----NIVPE-TSPESVPIISVSLPAGVIAARIGLWSFDLTVTQLL 480
Db 374 TISUSSALGVVFLHAVPYLVSP---PIAGRLVDTTGSYTAFLLCGFSMIFSSVL 426

RESULT 15

US-09-995-007-2
; Sequence 2, Application US/09995007
; Patent No. US20020123102A1
; GENERAL INFORMATION:
; APPLICANT: NANCY CARRASCO, ET AL.
; TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND
; NUCLEIC ACID ENCODING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE
; DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,007
; FILING DATE: 26-NO. US20020123102A1-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/595,553A
; FILING DATE: FEBRUARY 1, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG J. ARNOLD
; REGISTRATION NUMBER: 34,287
; REFERENCE/DOCKET NUMBER: 96700/393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELE: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: <Unknown>
; DESCRIPTION: PROTEIN
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: RAT

INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-007-2

Query Match 3.7%; Score 108; DB 10; Length 618;
Best Local Similarity 19.1%; Pred. No. 0.26;
Matches 103; Conservative 71; Mismatches 160; Indels 204; Gaps 27;

Qy 14 CGSLADYLTSAKFLLYLGHSLSTWGRDMHFVSVFVELYVGNLSLLLTAVYVGVVAGVQL 73
Db 132 CGTU-QYLVAT--WLYTG-----IYIYAPALILNQVTLGLDIWASLL 169
Qy 74 VLGAIIIGDWQKNAKLKVAQTSLVVQNVSVILCGIILMMVFLHKHELLTMYVHGW-LTSC 132
Db 170 STGIICITLYTTVGKMAVWTD--VFQVWVWLGVF-----WVILARG 209
Qy 133 YILIIITANTANLASTATAITIQ-----RDWIVVWAG----- 164
Db 210 VILGGPRNVLSLAQNHRSRLNLMDFDPDRSRYTFWTFIVGGTLVWLWSYGVNQAVQRY 269
Qy 165 ----EDRSKLANNNATIRRIDQLTNIIAPMAVGQIMTFGSPVIGCGFISGMNLSVMC 219
Db 270 VACHTEGAKLALL-----VNQL-----GLFLIVASAAC-CG----- 300
Qy 220 YVLLWKVYQK--TPALAVKAGLKEETEELKQLNLHKDTPEKP-----LEGTH 264
Db 301 --IVMFVYKDCDPLLTGRISAPDQYMPDLLVLDIFEDLPVPGFLACAYSGTILSTASTS 358
Qy 265 LMGVKDSNIHELEHEOEPTCASQMAEPRTFRDGVWYVYNQPVFLAGWGLAFLYMTVLGF 324
Db 359 INAMAATVEDLIKPRMPGLAPR-----KLVPIS-KGLSFIY---GS 396
Qy 325 DCITTGAVYTGGLSGSILSILMGASAITG---IMGTV-----AFT--WLRKCGL--VR 371
Db 397 ACLTV-----AALSSLLGGVQLGSGFTVMGVISGPLLAGFTLGMLLPACNTPGVL 446
Qy 372 TGLISGLAQLSCLILCVISVFMFGSPDLDSVFPEDIRSRFIQGESITPTKIPETITEY 431
Db 447 SGLAAGLA-VSLWVAVGATLYPPGE-----QTMGVLP TSAAC TND SV 488
Qy 432 M---SNGNSANIVPETSPE-----SVPIISVSLPAGVIAARIGLWSFDLTVTQLL 480
Db 489 LLGPPGATNASGNTGPPSSGMDTGRPALADTFYAISLYLYGALG-----TUTTML 536

Search completed: November 5, 2003, 18:38:42
Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:29:46 ; Search time 29 Seconds
(without alignments)
1893.528 Million cell updates/sec

Title: US-09-715-927-6
Perfect score: 2929
Sequence: 1 MTRAGDHNRRQGCCGLADY.....ACGPDAKEVRKENQANTSVV 571
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	612.5	20.9	524	T02499	hypothetical prote
2	564	19.3	498	T48385	transporter like p
3	503.5	17.4	551	B88949	protein R09B5.4 [i
4	209.5	7.2	677	C89010	protein R08F11.6 [
5	187.5	6.4	482	T01762	hypothetical prote
6	131.5	4.5	523	S06920	glucose transport
7	123.5	4.2	409	C96956	permease [imported
8	123.5	4.2	524	A31318	glucose transport
9	120.5	4.1	763	A31443	Na+/H+ antiporter
10	118.5	4.0	534	S14144	hexose transport p
11	118	4.0	437	A82140	hypothetical prote
12	117.5	4.0	783	E86254	hypothetical prote
13	116	4.0	469	C90446	permease, multidru
14	115.5	3.9	418	E83686	hypothetical prote
15	115	3.9	503	T26524	hypothetical prote
16	114	3.9	392	B95129	transporter, proba
17	114	3.9	392	G97999	hypothetical prote
18	113.5	3.9	470	A81804	transmembrane effl
19	113.5	3.9	482	B59803	metabolite transpo
20	113.5	3.9	522	A31556	glucose transport
21	113	3.9	435	A2079	hypothetical prote
22	112.5	3.8	435	S65982	transport protein
23	112	3.8	304	B83032	probable permease
24	112	3.8	461	D70073	metabolite transpo
25	112	3.8	1148	H90175	NADH dehydrogenase
26	111.5	3.8	403	E84063	multidrug resistan
27	111.5	3.8	473	G69789	sugar transporter
28	111.5	3.8	679	D69329	conserved hypothet
29	110	3.8	413	A87305	Na+/H+ antiporter

ALIGNMENTS

RESULT 1

T02499

hypothetical protein At2g38460 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein T19C21.5

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001

C;Accession: T02499; C84805

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, August 1998

A;Description: Arabidopsis thaliana chromosome II BAC T19C21 genomic sequence.

A;Reference number: Z14676

A;Accession: T02499

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-524 <ROU>

A;Cross-references: EMBL:AC004683; NID:g3395421; PID:g3395426

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayan, L.; Tallon, L.;

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84805

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-524 <STO>

A;Cross-references: GB:AE002093; NID:g3395426; PIDN:AAC28758.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g38460; T19C21.5

A;Map position: 2

A;Introns: 53/2; 169/2; 330/2; 387/3; 433/3; 470/3

Query Match 20.9%; Score 612.5; DB 2; Length 524;
Best Local Similarity 29.1%; Pred. No. 6e-40;
Matches 155; Conservative 95; Mismatches 179; Indels 103; Gaps 9;

Qy	28	LYGHSLSWTGDRMWHFVAVSVLVELYGNLSLLTAVYGLVAVGVVLGALIGDWVDKNA	87
Db	41	LYVGYFLARWSARTWEFSVALYMIHLWPNLSLLAAIYGAIESGSTAIFGPIVGQWEGMD	100
Qy	88	RLKVAQTSLVQVNSVILCG---IILMMVFLHKHLLTMVHGWLTSYIIILITANTAN	144
Db	101	YVKVRLRWLLFQNLSTYIAGGAVIKLLVSLDKSRNLPV-----FAILVLTNLG	151
Qy	145	----LASTATAITQRDWIVVVA-GEORSKLNANNATIRIDQLTNILAPMAVQIMTFG	199
Db	152	AIGVLSLTAGTILIERDWAVMVSEGHPPAVLTWNVIRGIDLSKLLSPVITGLIISFV	211
Qy	200	SPVIGCGFISGNWLVSMCVFVLLWKVYQKTPALAVKAGLKBETELKQLNLHKDTPKP	259
Db	212	SLKASAITFAAWATTAWVEWLFISVYGVPAIT-----RSNERRILRSRTKQVEGRDAP	267

Db 454 ITHIMQVITPERBNTVFGMHNALCQTSVLKDXLVILPLPATFAICIFISYGFVSCGH 513
QY 536 IMYFRF-----AONTLGNKLFACGPDKEVRKENQ 565
Db 514 MFFIYLVKTNLSIVGRKLSQL-PE-KKEEKEKE 546

RESULT 4
C89010
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89010
R:Anonymous, The C. elegans Sequencing Consortium.
S:Science 282, 2012-1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; PMID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89010
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-677 <STO>
A:Cross-references: GB:chr_V; PIDN:AAB54248.1; PID:G2088831; GSPDB:GN00023; CESP:R08F11.
A:Note: similar to mariner transposase
C:Genetics:
A:Gene: R08F11.6
A:Map position: 5

Query Match 7.2%; Score 209.5; DB 2; Length 677;
Best Local Similarity 26.1%; Pred. No. 2.5e-08;
Matches 65; Conservative 47; Mismatches 72; Indels 65; Gaps 11;

QY 135 LIITIANIANLASTATAITQORDWIVVAGEDRSKLANWNATIRIDQLTNILAPMAVGQ 194
Db 481 LAIFGAINRFLNAEKSMIARDVVALNG--KNLSRQNAWLTGFDQLNLISPIVVG 538
QY 195 IMT-FG--SPVIGCGFISGMNLVSMCVYVLLKVKYQKTPALAVKAGLKEETELKQLNL 251
Db 539 LLTFEGMHQTLIVFGI---YSMLSFLSKLIFLSLT- 571
QY 252 HKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRD---GWVS-YYNQ 306
Db 572 -----NQNV-SLQNRQEGYCSLPISDEEPLFKPKRPGVISTYWKQT 612
QY 307 VFLAGMGIAFLYMTVLGFDCTTGYAYTQGLSGSILSL-MGA-----SAITGIMGTVAFT 361
Db 613 TFCAAFGIALFYKTMGFDNLAVGYA----TSASDLSVITIGALKSYGAVAGWVGVI 668
QY 362 WLRRKCGLV 370
Db 669 FLEKRIHII 677

RESULT 5
T01762
hypochemical protein A_IG002P16.6 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01762
R:Miller, N.; Beck, C.; Kramer, J.
A:Note: submitted to the EMBL Data Library, June 1997
A:Reference number: Z14421
A:Cross-references: ENBL:AF007270; NID:G2191157; PID:G2191161; GSPDB:GN00063; ATSP:A_IG0
A:Description: The sequence of A. thaliana IG002P16.
A:Accession: T01762
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-482 <MIL>
A:Cross-references: ENBL:AF007270; NID:G2191157; PID:G2191161; GSPDB:GN00063; ATSP:A_IG0
C:Genetics:
A:Gene: ATSP:A_IG002P16.6
A:Map position: 5

A:Introns: 15/2; 86/3; 108/1; 179/3; 238/3; 263/3; 301/1; 396/2; 433/3

Query Match 6.4%; Score 187.5; DB 2; Length 482;
Best Local Similarity 20.2%; Pred. No. 8.5e-07;
Matches 84; Conservative 71; Mismatches 157; Indels 103; Gaps 14;

QY 26 FLLYLGHSLSTWGDMMHFAVSFLVELYGNLSLLTAVYGLVAGSVLVGLGAIIGDWVK 85
Db 107 YVLVASCCLVGNLVEQLMNFAPSAIMLY-PSLLPVAVMGFVTKLAIAGGPPVVGKFM 165
QY 86 NARLKVAQTSVLVQNVSVIL-CGIIIMVFLKHELLTMYHGWLVTSCYILLITIAN-IA 143
Db 166 SPRVPTYSILNVIQAAQVLSAGMI-----THAYTVPSTSSASSILLQPFALLPAGA 220
QY 144 NLAATATAITQORDWIVVAGEDRS-KLANWNATIRIDQLTNILAPMAVGQIMTFGSPV 202
Db 221 SLGFIASGVALERDWWLLAGINRPIALAAQNAVLRHLDL----- 261
QY 203 IGCGFISGMNLVSMCVYVLLKVKYQKTPAL--AVKAGLKEETELKQLNLHKDTEPK 260
Db 262 -----CETALIWLTKNFSSGVLDPRPKCSLNSCSAEGSRTN---TDSIVI 302
QY 261 EGTMLMGVKSNIHELEHEQEPTCASQMAEPRTFRDQWVSYYNQPVFLAGMGLAFL 320
Db 303 TKYHLL-----CSSVDIGMETIKLGWKEYIQQVLPASLAYVLLYFN 344
QY 321 VLGFDCITTYGYATQGLS-----GSIL-----SIIM-----GASA 350
Db 345 IV-----LTPGSLMTAFITQRCHLLRGLPLLFSSAWMKIVINLARVFLPTLSNYGL 400
QY 351 ITGIMGTVAFTWLRKRCGLVRLTGLISGLAQLSCLILCVISVFMPSGLDLSVSPF 405
Db 401 VMGVAATFLSANLVKRVGILKVKCFPAICYLE--LAGAVGLFFQASLLAVAVAVY 453

RESULT 6
S06920
glucose transport protein, hepatic - mouse
N:Alternate names: GLUT 2
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 24-Sep-1999
C:Accession: S06920; S05319; B44887; 148367
R:Suzue, K.; Lodish, H.F.; Thorens, B.
Nucleic Acids Res. 17, 10099, 1989
A:Title: Sequence of the mouse liver glucose transporter.
A:Reference number: S06920; MUID:90098776; PMID:2602116
A:Accession: S06920
A:Molecule type: mRNA
A:Residues: 1-523 <SUZ>
A:Cross-references: EMBL:X16986; NID:G55511; PIDN:CAA34855.1; PID:G55512
R:Asano, T.; Shibasaki, Y.; Lin, J.L.; Akanuma, Y.; Takaku, F.; Oka, Y.
Nucleic Acids Res. 17, 6386, 1989
A:Title: The nucleotide sequence of cDNA for a mouse liver-type glucose transporter prot
A:Reference number: S05319; MUID:89366666; PMID:2771649
A:Accession: S05319
A:Molecule type: mRNA
A:Residues: 1-105, 'D', 107-180, 'T', 182-326, 'T', 328-350, 'F', 352-431, 'P', 433-523 <ASA>
A:Cross-references: EMBL:X15684; NID:G51090; PIDN:CAA33719.1; PID:G51091
R:Hogan, A.; Heyner, S.; Charron, M.J.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Th
Development 113, 363-372, 1991
A:Title: Glucose transporter gene expression in early mouse embryos.
A:Reference number: A44887; MUID:92111400; PMID:1765007
A:Accession: B44887
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 384-431, 'P', 433-496 <HOG>
A:Cross-references: GB:S77926; NID:G242129; PIDN:AAB20847.1; PID:G242130
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIN:77926, NCBI:77927)
R:Waeber, G.; Thompson, N.; Haefliger, J.A.; Nicod, P.
J. Biol. Chem. 269, 26912-26919, 1994
A:Title: Characterization of the murine high Km glucose transporter GLUT2 gene and its t
A:Reference number: A55078; MUID:95014557; PMID:7929431

Db 214 NYDLWHILLGSGVRAILQSLLLFFCPSPRYLYIKL-----DEEVKAKQ-SLKR----- 262
Qy 261 EGTTHLMGVKD--SNIHELEHEOEPTCASQMAEPRTFRDGMVSYNQPVLGAM-GLAFL 317
Db 263 ----LRGVDDVTKOLINERKEREASQKVIITQLFTN---SSYRQDILVALMLHVAQQ 315
Qy 318 YMTVLGFCITTYGAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKRCGLVTRTGLISG 377
Db 316 FSGINGIFVYSTIFQTAGISKPVYATI-GVGAVNMV-FAVSVFLVEKAGRRSLFLGM 373
Qy 378 LAQLSCLI-----LCVISVFMFGSPDLDSVSPFDIRSRFIOGESITP 420
Db 374 SGMFCAIFMSVGLVLLNFKFMSYVSMIAIF-----LFVSFFE-----IGP 415
Qy 421 TKIPEITTEIYMSNGSNAVTPETSPESVPIIS-----PRPALAIAAFSNWTCNFVALCFQVIADFCGYVFFL 458
Db 416 GPIFWFVAEFSQ-----PRPALAIAAFSNWTCNFVALCFQVIADFCGYVFFL 469
Qy 459 FAGVIAA 465
Db 469 FAGVLLA 475

RESULT 9
AI3443
Na+/H+ antiporter 1 / proteinase IV (EC 3.4.-.-) [imported] - Brucella melitensis (strain Na+/H+ antiporter 1)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3443
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova, I.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:1175668
A:Accession: AI3443
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-763 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAU52716.1; PID:g17983545; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11535
A:Map position: I
C:Keywords: hydrolase

Query Match 4.1%; Score 120.5; DB 2; Length 763;
Best Local Similarity 21.5%; Pred. No. 0.26;
Matches 93; Conservative 55; Mismatches 150; Indels 135; Gaps 21;

Qy 55 GNSLLLTAVYGLVAGSV-----LVLGAI-IGDWVDKN-----ARLKVAQTS 95
Db 52 GITLMAAALALIVANSPFAQTYFDALHLYIGPLSLAHWINDALMAIFLLVGLLEIKREM 111
Qy 96 LVQNVQS-----VILCGIILMMVFLHKLHLLTMYHGWLTSCYIIILITIAN 141
Db 112 LDGQLASWPNRMLPGIAAGGVILPAIFAVL---NHDNPAKLRGWAVPSATDIAFALGV 168
Qy 142 IANLASTATAITQDRIWVVGEDRSKLANMNATIRRIQDLTNILAPMAVGOIMTFGSP 201
Db 169 LSLGSRAP-----SSLVFLATLAILDDLAAVI-IAIFYTAETISMP 210
Qy 202 VIGCGFTSG-----WNLVSMCEVYLLWKVYQKTPALAVKAGLKEEETELKQLN 250
Db 211 YLGAATIAAVLFVWNRMDVVKLLPYLISAVILMFFVFNKSGVHATVAGVAA----- 262
Qy 251 LHKDTEPKPLEGTHLMGVKD---SNIHELEHEOEPTCASQMAEPRTFRDGMVSYNQPV 307
Db 263 LMIFLKPAP-----GRPDDMTSPLKLEH-----ALAKP-----VAFIVVPI 299
Qy 308 F-LAGMGLAP--LYMTVLGFCITTYGAYTQGLSGSILSLMGASAITGIMGTVAFTWLR 364
Db 300 FGFANAGISFKGLEASVLG-DTLTLGI-----LLGLFLGKQ-----FGVFGAAWLA 344

Qy 365 RKQGLVTRTGLISGLAQL-SCLILCVISVFMFGSPDLDSVSPFDIRSRFIOGESITPTKI 423
Db 345 IKTLGAEKPMGASVQVLYGAILCGIGFTM-----SIFI-GLLSFPFSDL 387
Qy 424 PEITTEIYMSNGS 436
Db 388 MQTETKIGVLGS 400

RESULT 10
SI4144
hexose transport protein - Chlorella kessleri
C:Species: Chlorella kessleri
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1999
C:Accession: SI4144; S07096
R:Wolf, K.; Tanner, W.; Sauer, N.
Curr. Genet. 19, 215-219, 1991
A:Title: The Chlorella H(+)/hexose cotransporter gene.
A:Reference number: SI4144; MUID:91330335; PMID:1868571
A:Accession: SI4144
A:Molecule type: DNA
A:Residues: 1-534 <WOL>
A:Cross-references: EMBL:X55349; NID:g18040; PIDN:CAA39036.1; PID:g18041
R:Sauer, N.; Tanner, W.
FEBS Lett. 259, 43-46, 1989
A:Title: The hexose carrier from Chlorella. cDNA cloning of a eucaryotic H(+)-cotransporter.
A:Reference number: S07096; MUID:90092536; PMID:2599110
A:Accession: S07096
A:Molecule type: mRNA
A:Residues: 1-58,60-208,'L',210-247,'R',249-534 <SAU>
A:Cross-references: EMBL:Y07520; NID:g18038; PIDN:CAA68813.1; PID:g18039
C:Genetics:
A:Gene: HUD1
A:Introns: 29/3; 60/3; 79/2; 114/3; 152/1; 194/2; 234/2; 274/3; 309/1; 366/3; 405/3; 465/3; 465/3; 465/3
C:Superfamily: Glucose transport protein
C:Keywords: sugar transport; transmembrane protein
F:26-46/Domain: transmembrane #status predicted <TM1>
F:88-108/Domain: transmembrane #status predicted <TM2>
F:117-137/Domain: transmembrane #status predicted <TM3>
F:145-165/Domain: transmembrane #status predicted <TM4>
F:172-192/Domain: transmembrane #status predicted <TM5>
F:202-222/Domain: transmembrane #status predicted <TM6>
F:288-308/Domain: transmembrane #status predicted <TM7>
F:323-344/Domain: transmembrane #status predicted <TM8>
F:353-373/Domain: transmembrane #status predicted <TM9>
F:388-408/Domain: transmembrane #status predicted <TM10>
F:427-447/Domain: transmembrane #status predicted <TM11>
F:457-477/Domain: transmembrane #status predicted <TM12>

Query Match 4.0%; Score 118.5; DB 2; Length 534;
Best Local Similarity 19.4%; Pred. No. 0.24;
Matches 123; Conservative 90; Mismatches 213; Indels 207; Gaps 29;

Qy 5 GDHNRQRCGCSLADYLTGAKFLYLGLHSLSTWGRMWHFAVSVFLVE----- 52
Db 17 GDYR-----GGITVTVVWVAFWAAACGGLLLGYDNGVTGCVGSLEAFKKFFPDVWAKQ 70
Qy 53 -----LYGNSLLLTAVYGLVAGSVLVLAGIIGDWVDKNARLKVAQ----- 93
Db 71 EVHEDSPYCTYDNAKLQFLVSSSLFLAG---LVSCLFASMITRNWGRKVTWIGIGAPFVAG 127
Qy 94 --TSLVVQNVSVILCGIILM-----WVFLHKLHLLTMYHGWLTSCYIIILITIA-N 141
Db 128 GLVNFAQDMAMLIIVGVLLGVGVGLGSQVVPQYLSLEAFKKFFPDVWAKQ 187
Qy 142 IANLASTATAITQDRIWVVGEDRSKLANMNATIRRIQDLTNILAPMAV-GOIMTFGS 200
Db 188 IAGLVNVAV-----RW-----ENGWRLS-----LGPAAAPGAILFLGS 221
Qy 201 PVIGCGFISGWNILVSMCEVYLLWKVYQKTPALAVKAGLKEEETELKQLNKHDTPEKPL 260
Db 222 LVL-----PESPNFLVEKGTKEGREVLQ-----KL 247

Db 459 MSKFGFVLSLERRAARFYIFNLINVLGSGVITGSAFEQDSFL-----KOSAKEIP 513
Qy 444 ETPSPVPIISVSL-----FAGVIA--ARIGLWSFDLTWTQLQNVIESRGIING 494
Db 514 KTVGVAIKATFFITYTWGAGIAGEILRLKPLIFPHIKNSLVVTEKDRBAMMPG 573
Qy 495 VONSNNY-----LDDLHFMVILAPNPEAFGLLVLSVSVFAMGHIMY 538
Db 574 ---QINHYATBPRIQLYELGLGVAP---VTPVLLPFIIFALAYLVP 616

RESULT 13
C90446
permease, multidrug resistance protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 06-Jan-2003
C:Accession: C90446
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A39139
A:Accession: C90446
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <KUR>
A:Cross-references: GB:AB006641; NID:gl3816035; PIDN:AAK42826.1; GSPDB:GN00155
C:Genetics:
C:Gene: SSO2716
C:Superfamily: multidrug-efflux transporter

Query Match 4.0%; Score 116; DB 2; Length 469;
Best Local Similarity 22.7%; Pred. No. 0.32;
Matches 88; Conservative 64; Mismatches 143; Indels 92; Gaps 21;
Qy 68 VAGSVLVL--GAIIGDWDVKNARLVAQTSLVVQVNSVILGCIIMVMVFLHKHLLTWYH 125
Db 105 IGGSMNVANSSIVADVPPNRRKAGYIGTSLGNIGA-LVGIVLGV-----LITFF 156
Qy 126 GWLVTSCVILITITIANLANSTATAITIQDRIWVAGEDRSKLANMNAITRIDLTN 185
Db 157 GWQ-----YIFINV-----PIGI-----VAVILGV--MNIKDKINKVTKLDMGA 195
Qy 186 ILAPMAVQIMTEGSPVIGCGFISGNLVSMC-----VEYVLLMKVYQKTPALAVKAGLK 240
Db 196 ILLGLSLA-LISLSIMFTAASGISTDNIIEALGVALIPFFLLNETRSKYPILNKI--- 251
Qy 241 EETELKQLNLHKDTEPKPLEG-----THLMGVKDSNIHELEHEQEPTCASQMA 289
Db 252 -----FKIRLLTYSILANFLOGIGLSLSFLIMYLOQVR--GISPLDSSLITLTPGVIA 304
Qy 290 EPFTFRDGVWVYNNQPVFLAGMGLAFLYMT-VLGFDCITTTGYAYTOGLSGSILSMGA 348
Db 305 SILAPFM-GRVADRGKPGIVAGIGLMFTIFIMLYFFLLTPTDY-----YLIVGI 354
Qy 349 SAITGIMGTAVTWLRKCGLV-----RTGLISGLAQ-----LSCLILCVISVMPGSP 398
Db 355 SAITGI-GSAMP-WPSNSTAMFHAPKEYGVSGLSRLTGNIGTILSVLSIVVATLSI 412
Qy 399 DLSVSPPE-----DIRSRFTQG 415
Db 413 PRSVA-FEIFTGTTLNGDVSVTFVNG 438

RESULT 14
E83686
hypothetical protein BH0293 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E83686
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <STO>
A:Cross-references: GB:AP001508; GB:BA000004; NID:gi0172890; PIDN:BA04012.1; GSPDB:GN00155
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0293

Query Match 3.9%; Score 115.5; DB 2; Length 418;
Best Local Similarity 22.7%; Pred. No. 0.3;
Matches 112; Conservative 71; Mismatches 177; Indels 133; Gaps 27;
Qy 17 LADYLTSAKFL-----YLGHSLSWTGDRMWFHFAVSFVLVELYGNLSLLTAVYGLVVAG 70
Db 17 LSSLRNQKFLLLWVASIFSGLSFS-----IYFLTKSWYVVDLGLKTLFGLVLMAMTVP 71
Qy 71 SVLVGAIIGDWDVKNARLVAQTSLVVQVNSVILGCIIMVMVFLHKHLLTMYHGVLT 130
Db 72 RVLLM-MIGGVLDLRRFRSTIMPCSLTR--SVLLFGMFL-----QFLGVINIWTLI 121
Qy 131 SCYILIITI-----ANIANLASTATAITIQDRIWVVAGEDRSKLANMNAITRIDOLT 184
Db 122 GFAPCIGILDAPFPANVAILPS-----IVKKD-----HLTRANSFLOSTNHLT 165
Qy 185 NILAPMAVQIMTEGSPVIGCGFISGNLVSMC-----VEYVLLMKVYQKTPALAVKAGLK 240
Db 166 LIIGPMIAGWLTITTA-----YEIFVTIMCLLIIGAILIFRV--KEPF----- 207
Qy 241 EETELKQLNLHKDTEPKPLEGTHLMGVKQ--SNIHELEHEQEPTCASQMAEPFTPRDG 298
Db 208 -DKPETKRL-----PSPL-----KDLTEGYSVLRH--SPVLLSLM----- 239
Qy 299 WVSYYNQPVFLAGMGLAFLYMTV---LGFDCTTTGYAYTOGLSGSILSMGASAITGIM 355
Db 240 WMSVFVN-FTITPGTWAIPLIVVEERLKGDPLELSYLESTIAGLLIGAIM-----IGII 293
Qy 356 GTVAPTWLRKCGI-----VRTGLISGLAQSLCLILCVISVMPGSPDLSPVSPFEDI 408
Db 294 N-----LRSKRGLIIIRLTVMGVALLEGGVSLWQAIFVLAAILGWSIG-----DI 342
Qy 409 RSRFIQGESITPKIPITTEIYMSNGSNSANIVPETSPEVPI-----TSVSLIFAG-- 461
Db 343 PARAVVOEK---TDFDKIGRWGMGMSTASSGLIPLSYGMTSVALSLSLSSILLYCGSF 399
Qy 462 -VIAARIGLWSFD 473
Db 400 VILFSPFVLWKFD 412

RESULT 15
T26524
hypothetical protein Y18D10A.23 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26524
R;Harris, B.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z20226
A:Accession: T26524
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-503 <WIL>
A:Cross-references: EMBL:AL034393; PIDN:CAA22315.1; CESP:Y18D10A.23
A:Experimental source: clone Y18D10A
C:Genetics:
A:Gene: CESP:Y18D10A.23
A:Introns: 32/3; 98/3; 271/1; 340/1; 423/2; 441/1
Query Match 3.9%; Score 115; DB 2; Length 503;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:24:46 ; Search time 19 Seconds
(without alignments)
1413.276 Million cell updates/sec

Title: US-09-715-927-6

Perfect score: 2929

Sequence: 1 MTRAGDHNRRQCGGSLADY.....ACGPDKEVRKENQNTSW 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	126.5	4.3	523	1 GTR2 MOUSE	P14246 mus musculus
2	123.5	4.2	524	1 GTR2 HUMAN	P1168 homo sapien
3	118.5	4.0	534	1 HUPI1 CHLKE	P15686 chlorella k
4	113.5	3.9	482	1 YF1G BACSU	P54723 bacillus su
5	113.5	3.9	522	1 GTR2 RAT	P12336 rattus norv
6	112.5	3.8	435	1 YVBO BACSU	P37489 bacillus su
7	109.5	3.7	461	1 CSBC BACSU	P46333 bacillus su
8	108.5	3.7	400	1 RFAY_XANCP	P46358 xanthomonas
9	108	3.7	618	1 SL55 RAT	Q63008 rattus norv
10	106.5	3.6	492	1 YD19 METUA	Q58715 methanococc
11	106	3.6	547	1 YQJ1 ECOLI	P33941 escherichia
12	104.5	3.6	502	1 GTR5 RAT	P43427 rattus norv
13	103.5	3.5	455	1 AZAC_CAVPO	Q60476 cavia porce
14	103.5	3.5	533	1 GTR2_CHICK	Q90592 gallus gall
15	103	3.5	358	1 HH2R RAT	P25102 rattus norv
16	103	3.5	634	1 SL56 RAT	O70247 rattus norv
17	103	3.5	718	1 SL53 HUMAN	P53794 homo sapien
18	103	3.5	1043	1 P11D MOUSE	O35904 mus musculus
19	102.5	3.5	394	1 YWFF BACSU	P39642 bacillus su
20	102.5	3.5	429	1 ARSB_YEREN	P74985 yersinia en
21	102	3.5	547	1 MERA_STAP	P08663 staphylococ
22	101.5	3.5	458	1 AZAC MOUSE	Q01337 mus musculus
23	101.5	3.5	458	1 AZAC RAT	P22086 rattus norv
24	101.5	3.5	632	1 AFUB_HAEIN	Q57341 haemophilus
25	101.5	3.5	739	1 DID_HUMAN	P50443 homo sapien
26	101.5	3.5	971	1 Y277 MYCPN	P75387 mycoplasma
27	101	3.4	457	1 C1W2 ECOLI	P30868 escherichia
28	100.5	3.4	411	1 UIWB MOUSE	P97438 mus musculus
29	100.5	3.4	552	1 YH81 MYCTU	Q50684 mycobacteri
30	99.5	3.4	460	1 NU4M_SQUAC	Q92445 squalus aca
31	99.5	3.4	496	1 GTR3 HUMAN	P1169 homo sapien
32	98.5	3.4	472	1 EXUT ECOLI	P42609 escherichia
33	98.5	3.4	822	1 VGLE_HSV7J	P52352 human herpe

RESULT 1

ID	GTR2 MOUSE	STANDARD;	PRT;	523 AA.
AC	P14246; Q9DBA7;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Solute carrier family 2, facilitated glucose transporter, member 2			
DE	(Glucose transporter type 2, liver).			
GN	SLC2A2 OR GLUT2 OR GLUT-2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Liver;			
RX	MEDLINE=90098776; PubMed=2602116;			
RA	Suzue K., Lodish H.F., Thorens B.;			
RT	"Sequence of the mouse liver glucose transporter.";			
RL	Nucleic Acids Res. 17:10099-10099(1989).			
[2]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6; TISSUE=Liver;			
RX	MEDLINE=89366666; PubMed=2771649;			
RA	Asano T., Shibasaki Y., Lin J.L., Akanuma Y., Takaku F., Oka Y.;			
RT	"The nucleotide sequence of cDNA for a mouse liver-type glucose transporter protein.";			
RL	Nucleic Acids Res. 17:6386-6386(1989).			
[3]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Liver;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,			
RA	Schirml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA	Havashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
[4]				
RP	SEQUENCE OF 384-496 FROM N.A.			
RX	MEDLINE=92111400; PubMed=1765007;			

P18825 homo sapien
P03819 escherichia
Q48624 leuconostoc
P53793 bos taurus
P97046 lactococcus
P25336 saccharomyc
Q9184 rattus norv
Q12564 ampelomyces
P97292 mus musculus
O79881 sus scrofa
O87712 coxiella bu
Q9m3j4 spinacia ol

RA Hogan A., Heyner S., Charton M.J., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Thorens B., Schultz G.A.; Gilbert D.J.,
 RT "Glucose transporter gene expression in early mouse embryos."
 RL Development 113:363-372(1991).
 RN [5]
 RP TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6;
 RX MEDLINE=93170163; PubMed=1289053;
 RA Smith D.E., Gridley T.;
 RT "Differential screening of a PCR-generated mouse embryo cDNA library:
 RT glucose transporters are differentially expressed in early
 RL development 116:555-561(1992).
 CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
 CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
 CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
 CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
 CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
 CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
 CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
 CC YOLK SAC AND LIVER PRIMORDIUM.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTER SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X16986; CAA34855.1; -
 DR EMBL; X15684; CAA33719.1; -
 DR EMBL; S77926; AAB20847.1; -
 DR EMBL; AK005068; BAB23792.1; -
 DR PIR; S06920; S06920.
 DR MGD; MGI:1095438; Slc2a2.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; sugar_cr; 1.
 DR PRINTS; PRO0171; SUGRTNSPOT.
 DR TIGRFAMS; TIGR00879; SP; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Sugar transport; Transport; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 11 31 1 (POTENTIAL).
 FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 118 2 (POTENTIAL).
 FT DOMAIN 119 126 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 157 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 158 178 4 (POTENTIAL).
 FT DOMAIN 179 186 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 187 207 5 (POTENTIAL).
 FT DOMAIN 208 216 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 217 237 6 (POTENTIAL).
 FT DOMAIN 238 302 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 303 323 7 (POTENTIAL).
 FT DOMAIN 324 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 358 8 (POTENTIAL).
 FT DOMAIN 359 367 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 368 388 9 (POTENTIAL).
 FT DOMAIN 389 401 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 402 422 10 (POTENTIAL).
 FT DOMAIN 423 432 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 433 453 11 (POTENTIAL).

FT DOMAIN 454 460 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 461 481 12 (POTENTIAL).
 FT DOMAIN 482 523 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 106 106 G -> D (IN REF. 2).
 FT CONFLICT 181 181 T -> N (IN REF. 1).
 FT CONFLICT 327 327 S -> T (IN REF. 2).
 FT CONFLICT 351 351 F -> L (IN REF. 1).
 FT CONFLICT 432 432 P -> S (IN REF. 1).
 SQ SEQUENCE 523 AA; 57106 MW; F12B5376CDBD5F63 CRC64;
 Query Match 4.3%; Score 126.5; DB 1; Length 523;
 Best Local Similarity 22.3%; Pred. No. 0.11;
 Matches 107; Conservative 68; Mismatches 146; Indels 159; Gaps 27;
 QY 52 ELYGNSLLTAVGLVAGSVL--VLGATIGDWV-DKNARLKVQATSLVVQNVSVILCGI 108
 DB 84 ETEGSAHIVTMLSLSVSSFAVGGMVASFPGWGLGDKLRK-----AMLAAN-SLSLTGA 138
 QY 109 ILM--MVFLHKLHLLT-----MYHGWV--LTSCYVILITITIANIANLASTATAITQD 157
 DB 139 LLMGCKFGPAHALIAGRSVGLYGLISGLVPMY-----IGETAPT----- 182
 QY 158 WIVVAGEDRSKLANNNATIRRIDLTNILAPMAVGQIMTFGSPVIGCGFISG---WNL 213
 DB 183 -----LRGALGTLHQLALVTGILI-----SQIAGLSFILGNQDHHI 219
 QY 214 -----VSMCEVYLWKVYQKTPALAVKAGLKEETELKQLNLKDKTEKPLEGTHLMG 267
 DB 220 LLGLSAPALLQCLLLFCPSPRVLYIK--LEEVRKKSL-----KRLRGTEDEV- 268
 QY 268 VKDSNIHELHEHQEPTCASQMAEPPTFRDGVSVYVYNQPVFLAGM-GLAPLYMTVLGDFC 326
 DB 269 TKD--INEMKKKEKEASTEQKVSIOLEFTD---ANYRQPIVLVALMLHMAQQFSGINGIFY 323
 QY 327 ITTGAYVTOGLS-----GSI-----LSILMGASA-----ITGIMGTVAFTWLR 365
 DB 324 YSTISFQTAGISQPVYATIGVGNAINMFTAVSVLLVEKAGRRTLFLTGMIG-MFFCTIFM 382
 QY 366 KGLVNTGLISGLAQLSCLILCVISVFMFGSPDLDSVPFEDIRSRFIOGESITPTKPIE 425
 DB 383 SVGLV---LLDKFAWMSVYSMTAIFL-----VSFFE-----IGPGPIPW 419
 QY 426 ITTEIYVMSNGSNSANIVPETSPEVPIIS-----VSLLPAGVI 463
 DB 420 FMVAEPFSQG-----PRPTALAAFNWCNFWVIALCFQVIADFLGYPVFFLAGV 472
 RESULT 2
 GTR2_HUMAN STANDARD; PRT; 524 AA.
 AC P11168;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 2
 DE (Glucose transporter type 2, liver).
 GN SLC2A2 OR GLUT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT LEU-68.
 RC TISSUE=Kidney, and Liver;
 RX MEDLINE=88289735; PubMed=3399500;
 RA Fukumoto H., Seino S., Imura H., Sieno Y., Eddy R.L., Fukushima Y.,
 RA Byers M.G., Shows T.B., Bell G.I.;
 RT "Sequence, tissue distribution, and chromosomal localization of mRNA
 RT encoding a human glucose transporter-like protein."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:5434-5438(1988).
 CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
 CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA

```

FT          /FTid=VAR_014720.
SQ SEQUENCE      524 AA;   57489 MW;   DA60057720TEC083 CRC64;

Query Match              4.2%; Score 123.5; DB 1; Length 524;
Best local Similarity    20.3%; Pred.No. 0.18;
Matches                  99; Conservative                    150; Indels 169; Gaps 23;


Qy  41 MHFAVSVFLV-----ELYGNSL--LLTAVYGLVAAGSVLVLGAIIGDWVDKRNARLKVAQ 93
     :|||:|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   96 LWSLSYSFSAVGWMTASFFGGWLGDTLGRIKAMLVANILSLVGALLMGF-----SKLGP 149

Qy  94 TSLVV---QNVSIVLCGIIMVMPLKHHELLTMHGWLTSCVILIITIANIANTASTAT 150
     :::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db   150 SHILLTAGRSISGLYCGLI-----SGLVPMY-----IGEIAFTA- 183

Qy  151 AIIQRDIWWVAGEDRSKLANNMATIRRIDOLTNILAPMAVGQIMTFGPSVIICGFISG 210
     -----LRGALGTTHQLAIVTGILI-----SOIIGLEFILG 213

Qy  211 ----WNL-----VSMCEVEYLWKVKYTTPALAVALAKGLEKEETELKQLNLHKDTEPKPL 260
     ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   214 NYDLWHILLGSVRAILQSLLLFPFCPSPRLYIKL----DEEVKAQO-SLKR----- 262

Qy  261 EGHLMGVKD--SNIEHEHQEPTCASQAEBPFRTRDGWWSYYNQPVFLAGM-GLAFL 317
     ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   263 ---LRGYDVTKIDINEMKEREEASESEQKSVISIQLFNTN--SSYRQPILVALMLMHVAQQ 315

Qy  318 YMTVLGFDCTTCGYAYTOGLSGSILGISLGASAITGITMGTVAFTWLRRKCGLVRTGLISG 377
     ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   316 FSGINGIFYSTSIPOTAGISKRPVATI-VGVAVNWVF-TAVSVFLVEKAGRSLFLIGM 373

Qy  378 LAOLSCLI-----LCVISVFMFGSPLDLSVPFFEDIRSRFQGSEITP 420
     ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   374 SGMFEVCAlfMSvGLVLLNFkSMSyyVmIAIf-----LfVsffFe-----lGP 415

Qy  421 TKIPETTIIYMNGSNSANIvpETSpsvPIIs-----vsll 458
     ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   416 GPIEPWMVAEFESqg-----prPaalaiaafsnwntcnfiValCFQyiADfcgyvyvfFl 468

Qy  459 FAGVIAA 465
     |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   469 FAGULLA 475


RESULT 3
HUPl_CHLKE STANDARD; PRT; 534 AA.
AC PJ5686; Q39523;
DT 01-APR-1990 (Rel. 14, Created)
DD 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE H(+)/hexose cotransporter 1.
GN HUPl.
OS Chlorella kesslerii.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxId=3074;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90092536; PubMed=2599110;
RA Sauer N., Tanner W.;
RL "The hexose carrier from Chlorella. cdna cloning of a eucaryotic H+-cotransporter"; FEBS Lett. 259:43-46(1989).
[2]
SEQUENCE FROM N.A.
RP RP
RX MEDLINE=91330335; PubMed=1868571;
RA Wolf K., Tanner W.;
RL "The Chlorella H+/hexose cotransporter gene."; Curr. Genet. 19:215-219(1991).
CC -|- FUNCTION: ACTIVE UPTAKE OF HEXOSEs.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- INDUCTION: By glucose.
```

CC CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Y07520; CAA68813.1; -

CC EMBL; X55349; CAA39036.1; -

CC PIR; S14144; S14144.

CC InterPro; IPR007114; MFS.

CC InterPro; IPR005828; Sub_transporter.

CC InterPro; IPR005829; Sug_transporter.

CC InterPro; IPR003663; Sugar_transp.

CC Pfam; PF00083; sugar tr; 1.

CC PRINTS; PR00171; SUGRTRANSPORT.

CC TIGRfam; TIGR00879; SP; 1.

CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

CC Transmembrane; Transport; Sugar transport; Symport; Repeat;

CC Multigene family.

CC DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 22 42 1 (POTENTIAL).

CC DOMAIN 43 87 EXTRACELLULAR (POTENTIAL).

CC TRANSSEM 88 108 2 (POTENTIAL).

CC DOMAIN 109 114 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 115 135 3 (POTENTIAL).

CC DOMAIN 136 144 EXTRACELLULAR (POTENTIAL).

CC TRANSSEM 145 165 4 (POTENTIAL).

CC DOMAIN 166 173 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 174 194 5 (POTENTIAL).

CC DOMAIN 195 204 EXTRACELLULAR (POTENTIAL).

CC TRANSSEM 205 225 6 (POTENTIAL).

CC DOMAIN 226 299 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 300 322 7 (POTENTIAL).

CC DOMAIN 323 328 EXTRACELLULAR (POTENTIAL).

CC TRANSSEM 329 349 8 (POTENTIAL).

CC DOMAIN 350 352 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 353 373 9 (POTENTIAL).

CC DOMAIN 374 387 EXTRACELLULAR (POTENTIAL).

CC TRANSSEM 388 408 10 (POTENTIAL).

CC DOMAIN 409 433 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 434 454 11 (POTENTIAL).

CC DOMAIN 455 456 EXTRACELLULAR (POTENTIAL).

CC TRANSSEM 457 477 12 (POTENTIAL).

CC DOMAIN 478 534 CYTOPLASMIC (POTENTIAL).

CC TRANSSEM 60 60 MISSING (IN REF. 1).

CC CONFLICT 209 209 P -> L (IN REF. 1).

CC CONFLICT 248 248 C -> R (IN REF. 1).

CC SEQUENCE 534 AA; 57522 MW; 3B23B8E710AA2237 CRC64;

Query Match 4.0%; Score 118.5; DB 1; Length 534;

Best Local Similarity 19.4%; Pred.No. 0.41; Indels 207; Gaps 29;

Matches 123; Conservative 90; Mismatches 213; Indels 207; Gaps 29;

QY 5 GDNRRQCGCCGLADYLTSAKFLYLGHSLSTWGRMWHFAVSVFLVE----- 52

DB 17 GDYR-----GGLTVYVMVAFMAACGGLLGYDNGVTGGVSLFAFEKFFPDVWAKQ 70

QY 53 -----LYGNSLLLTATYGLVAGSVLVGAIGTDGVDKARLKVAG----- 93

DB 71 EVHEDSPCYTDYNAKLQFLVSSFLAG---LVSLFASWITRWGRKVTMGIGGAPFVAG 127

QY 94 --TSLVVQNVSVILCGIILM-----MVFLKHKELLTMYHGWLTSYVILITIA-N 141

DB 128 GLVNAFAQDMWMLVGRVLGFGVGLGSQVVPQVLSFVAFPSRGMNLNGYQLFVTIGIL 187

QY 142 IANLASTATAITIQDMVVVAVGEDSKLANMNATIRIDQLTNILAPNAV-QQIMTFGS 200

DB 188 IAGLVNAV-----RDW-----ENGWRLS-----LGPAAAPGAILFLGS 221

QY 201 PVIGCGFISGNLVSCVVEYVLLWKVYOKTPALAVKAGLKEBETELKOLNHLKDKPKPL 260

DB 222 LVL-----PSPNPLVEKGTKEGREVLQ-----KL 247

QY 261 EGTGLMGVKSNIHELEHQEPTCASQMAEPRTFRDGMVSYNQ---PVFLAGMGLAFL 317

DB 248 CGTSEVDAEFADI-----VAAVEIARPI-TMQSWASLFTRYMQLLTSTFVIQF- 296

QY 318 YMTVLGFDCCI-----TTGVAYTQGLSGSILSILMGASAITGIMGTVAFT-WLRRKC 367

DB 297 FQOFTGINAIFVVPVLFSSLSGANSAL---LNTVVGVAVNVGSTLIAMFSDKEGRRP 353

QY 368 GLVRTLISGLAQLSCLILCVISVMPGSG-PLDLSV-SPEDIRSRPIQGES----- 417

DB 354 LIEGGIOCLLAWLTGGVVLAEFAKYGTDPKAVASGILAVICIFISGFAMSGPMGW 413

QY 418 IPTPKIPEITTEIYMSNGSANSANIVPESVPIISVLSLLFAGVIAAIGLWFSFLT 477

DB 414 LIPSEI-----FILETTP-----AGTAVAVVGNFLFSFVIG 444

QY 478 QLLQENVIESERGIINGVQNSMNYLLDLHF-----IMVILAPNPFAFGLLVLSVFA 532

DB 445 QAFVSMLCAMEYGV-----FLFFAGWLIVMLCA-----IFLLPETKGVP 484

QY 533 MGHIMYFRAQNTLGNKLFACGPDAGEYRKENQ 565

DB 485 IERVOAL-YARHWFNRVM--GPAAAEVIAEDE 514

RESULT 4

YFIF_BACSU STANDARD; PRT; 482 AA.

AC P54723;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical metabolite transport protein yfif.

GN Yfif.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=168;

RC MEDLINE=96262713; PubMed=8704981;

RA Yamanoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;

RT "Determination of a 12 kb nucleotide sequence around the 76 degrees

RL Microbiology 142:1417-1421 (1996).

RN [2] SEQUENCE FROM N.A.

RP STRAIN=168;

RC MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Chou S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holtsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koeter P., Koningstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,


```

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandendool M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler H., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50543; BAA09111.1; -.
CC EMBL; Z99108; CAB12655.1; -.
CC PIR; B69803; B69803.
CC Subtilisin; BG11854; yfig.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGRTNSPORT.
CC TIGRfam; TIGR00879; SP; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT DOMAIN 1 29
FT TRANSMEM 30 50
FT DOMAIN 51 59
FT TRANSMEM 60 80
FT DOMAIN 81 92
FT TRANSMEM 93 113
FT DOMAIN 114 120
FT TRANSMEM 121 141
FT DOMAIN 142 155
FT TRANSMEM 156 176
FT DOMAIN 177 184
FT TRANSMEM 185 205
FT DOMAIN 206 263
FT TRANSMEM 264 284
FT DOMAIN 285 301
FT TRANSMEM 302 322
FT DOMAIN 323 331
FT TRANSMEM 332 352
FT TRANSMEM 353 373
FT DOMAIN 374 400
FT TRANSMEM 401 421
FT DOMAIN 422 423
FT TRANSMEM 424 444
FT DOMAIN 445 482
FT TRANSMEM 482 482
SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48EE01 CRC64;
Query Match 3.9%; Score 113.5; DB 1; Length 482;
Best Local Similarity 19.7%; Pred. No. 0.83;
Matches 77; Conservative 73; Mismatches 154; Indels 87; Gaps 19;
QY 34 LSTGDRMHWFAVSVFLVEL----YGNLLITAVYGLVAGSVLV---LGLAIGDWV-DK 85
DB 27 VSTGGLLFGYDTGVINGALFPMATAGQNLPTPTGVLVASSLLGAAFGAFMFGRLSDR 86
QY 86 NARLKVAQTSILVQNVSVILCGIILMMVFLHKLHLLTMVHGWLVTSCVILIITIANL 145

```

```

Db 87 HGRRK-----TILYLALLFIAATL-----GCTSPNASVMIAFRFLGL 125
QY 146 ASTATAITIQEDWIVVAGEDRSKLANNATIRIDOLTNILAPWAVQINTFSPVIGC 205
Db 126 AVGCASVTPTFLAISPARRGRIVTONELMIVIGQLLAYTFNAIIGSTMGESANV--- 182
QY 206 GFISGMN--LVSMCVYVLLW---KVYQKTPA-LAVKAGLKEEBTELKQLNLHDKTEPKP 259
Db 183 -----WYMLVIATLPAVWLFVFGMLIVPESPRWLAAGRMDALVLRQ--IREDSQAQ- 234
QY 260 LEGTHLMGVKDSNIHELEHEQEPTCASQMAEPPTFRDGVSVYVQVFLVAGMGLAFLYM 319
Db 235 -----QETKEIKHAIEGTAKK---AGHFDFQEPMI---RRILFIGIGIA-IVQ 275
QY 320 TVLGFDCI-----TTGVAYTQGLSGSILSLMGASAITGIMGTVAFTWL---RRK 366
Db 276 QITGVNSIMYGTETILREAGFQTEAALIGMI-----ANGVISVIAVIFGILWLGKVR 329
QY 367 CGLVRTGLISGLAQLSCLILCVISVFMFGSP 397
Db 330 PMLI-IGQIGTMTAL--LLIGILSIVLEGTP 357

RESULT 5
GTR2 RAT
ID GTR2 RAT STANDARD; PRT; 522 AA.
AC P12336;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 2
DE (Glucose transporter type 2, liver).
GN SLC2A2 OR GLUT2 OR GLUT-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=89003066; PubMed=3048704;
RA Thorens B., Sarkar H.K., Kaback H.R., Lodish H.F.;
RT "Cloning and functional expression in bacteria of a novel glucose
RT transporter present in liver, intestine, kidney, and beta-pancreatic
RT islet cells."
RL Cell 55:281-290(1988).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND
CC BETA-PANCREATIC ISLET CELLS.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03145; AAA41298.1; -.
CC PIR; A31556; A31556.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.

```

```

DR PRINTS: PR00171; SUGRTRANSORT.
DR TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
DR DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 11 31 1 (POTENTIAL).
FT TRANSMEM 11 31 1 (POTENTIAL).
FT DOMAIN 32 96 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 97 117 2 (POTENTIAL).
FT DOMAIN 118 125 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 126 146 3 (POTENTIAL).
FT DOMAIN 147 156 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 157 177 4 (POTENTIAL).
FT DOMAIN 178 185 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 186 206 5 (POTENTIAL).
FT DOMAIN 207 215 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 216 236 6 (POTENTIAL).
FT DOMAIN 237 301 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 302 322 7 (POTENTIAL).
FT DOMAIN 323 337 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 338 358 8 (POTENTIAL).
FT DOMAIN 359 365 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 366 386 9 (POTENTIAL).
FT DOMAIN 387 401 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 402 422 10 (POTENTIAL).
FT DOMAIN 423 431 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 432 452 11 (POTENTIAL).
FT DOMAIN 453 459 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 460 480 12 (POTENTIAL).
FT DOMAIN 481 522 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 522 AA; 57085 MW; 075AB81B56CF33F7 CRC64;

Query Match 3.9%; Score 113.5; DB 1; Length 522;
Best Local Similarity 21.7%; Pred. No. 0.91;
Matches 102; Conservative 65; Mismatches 164; Indels 139; Gaps 26;

QY 52 ELYGNLLLTAVGLVAGSVL--VLGATIGDWV-DKNARKLVAQSLVQNVSVILCGI 108
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 ETEGSAHIVTMLSLSVSPFVAGVWSPFGLGDKLGRK-----AMLAAN-SLSLTGA 137
QY 109 ILM--NVFLHKLHLLT-----MYHGWV--LTSVILLIITIANIANLASTATAITQRD 157
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 138 LLMGCKSGFPAHALIITAGRSVSGLYCLISGLVPMY-----IGEIAPT----- 181
QY 158 WIVVWAGEDRSKLANMNATIRIDQLTNILAPMAVGQIMTFGSPVIGCGFISG----WNL 213
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 -----LRGALGTLHLALVTGILI-----SQIAGLSFILGNQDYWHI 218
QY 214 -----VSMCVFVLLWKVYQKTPALAVKAGLKEEETELKQLNLHKDTEPKPLEGTHLMG 267
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 219 LLGLSAPVALLQCLLL--LFCPSPRLYLNLLEEVRACKSL-----KRLRGTEDI- 267
QY 268 VKDSNTHLEHEOEPTCSQMAQBPFRFDGWVSYNQPVFLAGM--GLAFLYMTVLGFDC 326
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 268 TKD--INEMKEKEASTEQKSVIQLFTD---PNYRQPIVVALMLHLAQFSGINGIFY 322
QY 327 ITTCYAYTQGLSGSIILMGASAITGIMGTVAFTLRRKCGIVRTGLISG-LAQSLCLI 385
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 323 YSTSIQTAGISQPVYATI-GVGAINMIF-TAVSVLLVERAGR-RTLFLAGMIGMFFCAV 379
QY 386 LCVISV-----FMPGSPDLDSVPPEDIRSRFIQESITPKPIEITTEIYWSNG 435
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 380 FMSLGLVLLDKFTWMSVYSMTAIFLFSVFFE-----ICPGPIFMVMAEFPFSG 428
QY 436 SNSANIVPETSPPSPVPIIS-----VSLLPAGVI 463
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 429 -----PRATALAALAFSNWVCNFIALCFQVIADFLGYPVFLFAGV 471

RESULT 6
YBBO_BACSU

```

```

ID YBBO_BACSU STANDARD; PRT; 435 AA.
AC E37489;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical transport protein ybbo.
GN YBBO.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsch M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Chou S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.F.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue J.C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra F., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE PHTHALATE PERMEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D26185; BAA05188.1; -
CC EMBL; Z99124; CAB16094.1; -
CC PIR; S65982; S65982.
CC Subtilist; BG10016; ybbo.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
KW TRANSMEM 26 46 POTENTIAL.
FT

```

FT	TRANSMEM	61	81	POTENTIAL.
FT	TRANSMEM	96	116	POTENTIAL.
FT	TRANSMEM	119	139	POTENTIAL.
FT	TRANSMEM	150	170	POTENTIAL.
FT	TRANSMEM	177	197	POTENTIAL.
FT	TRANSMEM	242	262	POTENTIAL.
FT	TRANSMEM	281	301	POTENTIAL.
FT	TRANSMEM	325	345	POTENTIAL.
FT	TRANSMEM	347	367	POTENTIAL.
FT	TRANSMEM	385	405	POTENTIAL.
FT	TRANSMEM	407	427	POTENTIAL.
SQ	SEQUENCE	435 AA;	48248 MW;	9630914D9B606208 CRC64;

Query Match 3.8%; Score 112.5; DB 1; Length 435;
Best Local Similarity 20.2%; Pred. No. 0.87;
Matches 97; Conservative 74; Mismatches 186; Indels 123; Gaps 20;

QY	20	YLTSKFLLYLGHSLSTWGDWMHFAVSVLVELVYGNLSLLTAVYGLVA	---GSVLVLG 76
DB	22	YISS-----LLSGITLNTFDRV---AISVAIPAQDSFHLTATELGVFSIVTSYTLMQ	74
QY	77	AIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKHELITMHWGLTSCYILI	136
DB	75	LPVGSLLDRFGVAWTVRVGMTIWSFLTIL-----LAFLOGKLLLYLFR	121
QY	137	ITIANIANLASTAITIQRDWVIVVAGEDRSKLANNNATIRRIDQLTNIL-APMAVGQI	195
DB	122	LTSASAFPAASKATAL-----WF-----PPSERGLANSLFDSAAKFSNVIGAPLVAFLV	170
QY	196	MTEGSPVIGCGFISGNNLVSM-----CVEY---VLLWKVYQKTPALAVKAGLKEETELKQL	249
DB	171	TTF-----DWRVAFUTIGCINVLFTIFFWQYYEQP-----ERHKRISKSELNYI	214
QY	250	NLHKDTEPKPLEGTHLMGVKDSNIHELEHGEQPTCASOMABPFRTRDGVWSYYNQPVFL	309
DB	215	QKH-----NAITEQIPYKTGPLKKLFTNRKVV	243
QY	310	AGMGLAFLYMTVLGDFCITTYAYTQGLSGSILSMGASAITGIMGTVAFT--- <td>366</td>	366
DB	244	--GL-----MIGP---TGCVTFNLLTLWLPFFKHKTGYMDLMSGLFTAVPWL---	287
QY	367	CGLVRTGLISGLAQLSCILLCVISVFMGSPGLDLSVSPEDIRSRFQIGESITPT-KIPE	425
DB	288	-----ISTSIGAVGGVLVDYFIKGYENTKYRTVIIVGMSFGFFFLGSLTNNITVAI	342
QY	426	ITTEIYMSGNSANIVPETSPESPIISVSL-----LFAGVIAARIGLWSPDLT	475
DB	343	ICISIGLAGIATAPVGHISIAELAPICGSVMSLWSWNLANNLFGGIIPASLTGYLFDVT	402

RESULT 7
CSBC_BACSU STANDARD; PRT; 461 AA.
AC P46333; O32289;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable metabolite transport protein cabc.
GN CSBC OR S592BR.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=1423;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1AL;
RX MEDLINE=96093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RL genome between the gnt and iol operons.";
RN DNA Res. 2:61-69(1995).
RE [2]
RP REVISIONS.
RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.

FT	TRANSMEM	39	59	2 (POTENTIAL).	RC	STRAIN=ATCC 33913 / NCPPB 528;
FT	DOMAIN	60	76	CYTOPLASMIC (POTENTIAL).	RX	MEDLINE=22022145; PubMed=12024217;
FT	TRANSMEM	77	97	3 (POTENTIAL).	RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
FT	DOMAIN	98	104	EXTRACELLULAR (POTENTIAL).	RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
FT	TRANSMEM	105	125	4 (POTENTIAL).	RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
FT	DOMAIN	126	139	CYTOPLASMIC (POTENTIAL).	RA	Camarotte G., Cannaval F., Cardozo J., Chamberg F., Ciapina L.P.,
FT	TRANSMEM	140	160	5 (POTENTIAL).	RA	Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
FT	DOMAIN	161	163	EXTRACELLULAR (POTENTIAL).	RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
FT	TRANSMEM	164	184	6 (POTENTIAL).	RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
FT	DOMAIN	185	241	CYTOPLASMIC (POTENTIAL).	RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
FT	TRANSMEM	242	262	7 (POTENTIAL).	RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
FT	DOMAIN	263	280	EXTRACELLULAR (POTENTIAL).	RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
FT	TRANSMEM	281	301	8 (POTENTIAL).	RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
FT	DOMAIN	302	308	CYTOPLASMIC (POTENTIAL).	RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
FT	TRANSMEM	309	329	9 (POTENTIAL).	RA	Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
FT	DOMAIN	330	341	EXTRACELLULAR (POTENTIAL).	RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
FT	TRANSMEM	342	362	10 (POTENTIAL).	RA	Setubal J.C., Kitajima J.P.;
FT	DOMAIN	363	378	CYTOPLASMIC (POTENTIAL).	RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
FT	TRANSMEM	379	399	11 (POTENTIAL).	RT	host specificities.";
FT	DOMAIN	400	402	EXTRACELLULAR (POTENTIAL).	RL	Nature 417:459-463(2002).
FT	TRANSMEM	403	423	12 (POTENTIAL).	RN	SEQUENCE OF 1-207 FROM N.A.
FT	DOMAIN	424	461	CYTOPLASMIC (POTENTIAL).	RP	MEDLINE=96025094; PubMed=7579621;
FT	CONFLICT	400	401	RP -> SA (IN REF. 3).	RX	Dow J.M., Osbourn A.E., Wilson T.J., Daniels M.J.;
SQ	SEQUENCE	461 AA; 50235 MW; B6587B5DC7272EF1 CRC64;			RA	"A locus determining pathogenicity of Xanthomonas campestris is
					RT	involved in lipopolysaccharide biosynthesis.";
					RL	Mol. Plant Microbe Interact. 8:768-777(1995).
					CC	-1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
					CC	ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
					CC	THEN IS RELEASED. THIS SIGMA FACTOR IS INVOLVED IN
					CC	LIPOLYSACCHARIDE BIOSYNTHESIS AND PATHOGENICITY.
					CC	-1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY. ECF SUBFAMILY.
					CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
					CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
					CC	the European Bioinformatics Institute. There are no restrictions on its
					CC	use by non-profit institutions as long as its content is in no way
					CC	modified and this statement is not removed. Usage by and for commercial
					CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/
					CC	or send an email to license@isb-sib.ch).
					CC	EMBL; AB012213; AAM40442.1; -
					DR	EMBL; U19896; AAA92044.1; -
					DR	InterPro: IPR000838; Sigma70_ECF.
					DR	Pfam: PF04542; sigma70_22; 1.
					DR	PROSITE: PS01063; SIGMA70_ECF; 1.
					KW	Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
					KW	DNA-binding; Complete proteome.
					FT	DOMAIN 62 75 POLYMERASE CORE BINDING (POTENTIAL).
					FT	DNA BIND 165 184 H-T-H MOTIF (BY SIMILARITY).
					FT	CONFLICT 1 6 MHADTL -> MLVPGHFRCTPTPW (IN REF. 2).
					FT	CONFLICT 19 19 A -> G (IN REF. 2).
					FT	CONFLICT 23 23 Q -> K (IN REF. 2).
					FT	CONFLICT 31 31 A -> V (IN REF. 2).
					FT	CONFLICT 38 38 A -> P (IN REF. 2).
					FT	CONFLICT 57 57 A -> V (IN REF. 2).
					FT	CONFLICT 69 69 H -> Y (IN REF. 2).
					FT	CONFLICT 80 80 S -> E (IN REF. 2).
					FT	CONFLICT 99 99 S -> I (IN REF. 2).
					FT	CONFLICT 106 106 A -> P (IN REF. 2).
					FT	CONFLICT 168 168 A -> G (IN REF. 2).
					FT	CONFLICT 200 200 A -> V (IN REF. 2).
					FT	CONFLICT 206 207 TA -> DR (IN REF. 2).
					SQ	SEQUENCE 400 AA; 42322 MW; 4C19F5E15CE91FAE CRC64;
						Query Match 3.7%; Score 108.5; DB 1; Length 400;
						Best Local Similarity 21.4%; Pred. No. 1.5;
						Matches 77; Conservative 52; Mismatches 116; Indels 115; Gaps 17;
QY					115	LHKHEL-----LTMVHGWLTSVILLITIANIANLASTATAITQRIWVWVAGEDRS 168
DB					10	LH-HELPHAAAGCQQAQYGRIVTACQNTV-----TATAIATRD---VAASEDIA 54

```

RESULT 9
SL55_RAT      STANDARD;      PRT;      618 AA.
ID            SL55_RAT
AC            Q63008;
DT            16-OCT-2001 (Rel. 40, Created)
DT            16-OCT-2001 (Rel. 40, Last sequence update)
DT            28-FEB-2003 (Rel. 41, Last annotation update)
DE            Sodium/iodide cotransporter (Na+/I-) cotransporter) (Sodium-iodide
DE            symporter) (Na+/I--symporter).
DE            SLC5A5 OR NIS.
OS            Rattus norvegicus (Rat).
OC            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX            NCBI_TaxID=10116;
RN            [1]
SEQUENCE FROM N.A.
RC            TISSUE=Thyroid;
RX            MEDLINE=96158880; PubMed=8559252;
RT            Dai G., Levy O., Carrasco N.;
RA            "Cloning and characterization of the thyroid iodide transporter.";
RL            Nature 379:458-460(1996).
CC            N1 - FUNCTION: MEDIATES IODIDE UPTAKE IN THE THYROID GLAND.
CC            -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC            -1- SIMILARITY: BELONGS TO THE SODIUM:SOLUTE SYMPORTER FAMILY (SSF).
CC            -----
CC            This SWISS-PROT entry is copyright. It is produced through a collaboration
CC            between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC            the European Bioinformatics Institute. There are no restrictions on its
CC            use by non-profit institutions as long as its content is in no way
CC            modified and this statement is not removed. Usage by and for commercial
CC            entities requires a license agreement (See http://www.iesb-sib.ch/announce/
CC            or send an email to license@sib-sib.ch).
CC            -----
DR            EMBL; U60282; AAB03338.1; -.
DR            PIR; S68513; S68513.
DR            InterPro; IPR001734; Na/solut_symport.
DR            Pfam; PF00474; SSF; 1.
DR            TIGRfams; TIGR00813; sss; 1.
DR            PROSITE; PS00456; NA_SOLUT_SYMP_1; 1.
DR            PROSITE; PS00457; NA_SOLUT_SYMP_2; FALSE_NEG.
DR            PROSITE; PS50283; NA_SOLUT_SYMP_3; 1.
DR            Transport; Transmembrane; Sodium transport; Symport; Glycoprotein;
DR            Phosphorylation.
KT            DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT            TRANSMEM 17 37 POTENTIAL.
FT            DOMAIN 38 53 CYTOPLASMIC (POTENTIAL).
FT            TRANSMEM 54 74 POTENTIAL.
FT            DOMAIN 75 88 EXTRACELLULAR (POTENTIAL).
FT            TRANSMEM 89 109 POTENTIAL.
FT            DOMAIN 110 136 CYTOPLASMIC (POTENTIAL).
FT            TRANSMEM 137 157 POTENTIAL.
FT            DOMAIN 158 163 EXTRACELLULAR (POTENTIAL).

```


RESULT 12

```
FT TRANSMEM 132 152 4 (POTENTIAL).
FT DOMAIN 153 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 181 5 (POTENTIAL).
FT DOMAIN 182 191 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 192 212 6 (POTENTIAL).
FT DOMAIN 213 276 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 277 297 7 (POTENTIAL).
FT DOMAIN 298 314 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 315 335 8 (POTENTIAL).
FT DOMAIN 336 342 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 343 363 9 (POTENTIAL).
FT DOMAIN 364 367 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 368 388 10 (POTENTIAL).
FT DOMAIN 389 411 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 412 432 11 (POTENTIAL).
FT DOMAIN 433 438 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 439 459 12 (POTENTIAL).
FT DOMAIN 460 502 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 164 164 V > A (IN REF. 1).
FT CONFLICT 285 285 A > T (IN REF. 2).
FT CONFLICT 452 452 T > S (IN REF. 2).
FT CONFLICT 479 479 K > N (IN REF. 1).
SQ SEQUENCE 502 AA; 55543 MW; 871FEDEFA1E7894D CRC64;

Query Match
Best Local Similarity 3.6%; Score 104.5; DB 1; Length 502;
Matches 121; Conservative 88; Mismatches 211; Indels 185; Gaps 31;

QY 15 GSLADYLTSAKFLYLGHSLSTWGRMHPFAVSFLVLYLXGN-----SLLLTAV 63
DB 10 GKLTVLALATFLAAGSSFOYGVNVAVNSPSEFMQOYFNDYVDRNKENIESFTLTLL 69
QY 64 YGLVVA-----GSVLVGAIGLDWVDKARL-----KVAQT--SLV 97
DB 70 WSLTVSFPFGPGFISLMV-GFLVNNLGRKALLFNFIISILPAILMGCSKIAKSPFII 128
QY 98 VQNVSVILC-GIILMVFHLKHLLTMYHGWLTSCVILLIITA-NIANLASTATAITIQ 155
DB 129 ASRLLVGICAGISNVVPMVLGELAPKNLRCALGVVLPOLFITVGLVLAQLFGLRSVLASE 188
QY 156 RDMIVVAGEDRSKLANMNATIRIDQTLNIPAMVGOIMTF-GSPVIGCGFISGNLV 214
DB 189 EGNPILIG-----LTGVPAGLQLLLLPFFPESP----- 216
QY 215 SMCVEVLLMKVYQKTPALAVKAGLKEEETELQNLNKHDKTEPKPLEGTHLMGVKDSNIH 274
DB 217 -----RYLLIQK--KNESAIEKA-----LQTLRGWKDV-----DMEME 247
QY 275 ELEHEQBPCTCASQMAEPFRIFRD---GWVSYYNOPVFLAGMGLAFLYMTVLGFDCLITTY 331
DB 248 EIRKEDEAEKAGFISVWKLFPMQSLRW-QLISTIVLMAGQQLS-----GVNAI---Y 296
QY 332 AYTOGL-----SGSILSILMGASAITGIMGVAF-----TWLRRKGLVRTGLISGLAQ 380
DB 297 YYADQIVLSAGKSNVDVQVYTAGTGA VNVFWMTVTVFVVELWGRNL-----LLIGFS- 349
QY 381 LSCILICVISVFMGSPGLDSVSPFEDIRS--RFIQGESITPKIPEI-TTEIYMGNSN 437
DB 350 -TCLTACIVLTVALALQNTISWMPYVSVICVIVVIVGHAVGPSIFALFITEIFLOSSRP 408
QY 438 SANIVPETSPESPYIIS--VSLLFAGVIAARIGLMSFDLTVTQLQENVIIESERGIING 494
DB 409 SAYMI-----GGSVHWLSNFIUGLIFPFI-----QVGLGPYFIIIFAI----- 447
QY 495 VQNSMNYLLDLHLHFIMVILAPNEAFGLLLVLSVFMAGHIMYFRPAQNTLGNKLFACG 554
DB 448 -----CLLTIVIFVMV-----PETGR-----TFVEINQI-----FAKK--NKVSDVY 484
QY 555 PDAKE 559
DB 485 PEKEE 489
```

```
RESULT 13
A2AC_CAVPO STANDARD; PRT; 455 AA.
ID A2AC_CAVPO
AC Q60476;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-2C adrenergic receptor (Alpha-2C adrenoceptor).
GN ADRA2C.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley;
RA MEDLINE=96152573; PubMed=8573196;
RX Svensson S.P., Bailey T.J., Porter A.C., Richman J.G., Regan J.W.;
RT "Heterologous expression of the cloned guinea pig alpha 2A, alpha 2B,
RT and alpha 2C adrenoceptor subtypes. Radioligand binding and
RT functional coupling to a CAMP-responsive reporter gene.";
RL Biochem. Pharmacol. 51:291-300(1996).
CC -!- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC PROTEINS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; U25724; AAA67076.1; -
CC HSSP; P29274; LMMH.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm1; 1.
CC PRINTS; PR00237; GPCR_RHODOPSIN.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
CC PROSITE; PS00362; G-PROTEIN RECEPTOR FL2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family; Phosphorylation.
CC DOMAIN 1 47 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 48 72 1 (POTENTIAL).
CC DOMAIN 73 84 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 85 110 2 (POTENTIAL).
CC DOMAIN 111 120 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 121 143 3 (POTENTIAL).
CC DOMAIN 144 164 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 165 187 4 (POTENTIAL).
CC DOMAIN 188 203 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 204 227 5 (POTENTIAL).
CC DOMAIN 228 376 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 377 400 6 (POTENTIAL).
CC DOMAIN 401 413 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 414 434 7 (POTENTIAL).
CC DOMAIN 435 455 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT DISULFID 120 198 BY SIMILARITY.
CC SQ SEQUENCE 455 AA; 6B657D247FF8A3F4 CRC64;

Query Match
Best Local Similarity 3.5%; Score 103.5; DB 1; Length 455;
Matches 54; Conservative 36; Mismatches 80; Indels 49; Gaps 11;

QY 60 LTAVYGLV-----VAGSVLVGAIIGDWVDKARLKVQATSLVVQNVSVILCGIILMMVFL 115
DB 50 LAADVGFLLVFTVGVNVLVVIIVL-----TSRALRAPQNLFLVLSASADILVATLVMPFS 104
```


QY 116 HKHELLT-MYHGWLTSCVI---LIITANTANLASTATAITQDRIWVWAGEDRSKLA 171
Db 105 LANELMAYWYFQVWGVYALDVLFCSTSSIVHLC---AISLDYWSVTQAVE-----Y 155
QY 172 NNATIRRIDQLTNLAPWAVQIMTF-----GSPVIGGPF-ISGNWLVSMCV- 218
Db 156 NLKTRPRVK--ATIVAVWLISAIISFPLVSFYRQPDGAAYPRGNDLWYILSSCIG 213
QY 219 -----EYVLMLKVQY-KTPALAVKAGLKEETE 245
Db 214 SFFAPCLIMGLVYRIYVAKLRTLTSEKRGPAPEGE 252
RESULT 14
GTR2_CHICK
ID_GTR2_CHICK STANDARD; PRT; 533 AA.
AC Q90592;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 2
DE (Glucose transporter type 2, liver).
GN SLC2A2 OR GLUT2 OR GLUT-2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94213485; PubMed=8161202;
RA Wang M.Y., Tsai M.Y., Wang C.;
RT "Identification of chicken liver glucose transporter.";
RL Arch. Biochem. Biophys. 310:172-179(1994).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOPFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCYTOLOGICAL TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z22932; CAA80519.1; -;
DR PIR; S43230;
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 18 38 1 (POTENTIAL).
FT DOMAIN 39 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 131 2 (POTENTIAL).
FT DOMAIN 132 136 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 137 157 3 (POTENTIAL).

FT DOMAIN 158 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 184 4 (POTENTIAL).
FT DOMAIN 185 199 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 200 220 5 (POTENTIAL).
FT DOMAIN 221 229 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT DOMAIN 251 315 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 316 336 7 (POTENTIAL).
FT DOMAIN 337 350 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 351 371 8 (POTENTIAL).
FT DOMAIN 372 379 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 380 400 9 (POTENTIAL).
FT DOMAIN 401 413 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 414 434 10 (POTENTIAL).
FT DOMAIN 435 445 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 446 466 11 (POTENTIAL).
FT DOMAIN 467 471 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 472 492 12 (POTENTIAL).
FT DOMAIN 493 533 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 533 AA; 57699 MW; D06284CDC2779A3D CRC64;
Query Match 3.5%; Score 103.5; DB 1; Length 533;
Best Local Similarity 20.2%; Pred. No. 4.9;
Matches 89; Conservative 70; Mismatches 129; Indels 153; Gaps 23;
QY 42 WHEAVSVFLVELYGNLSLLTAVYGLVWAGSVLVGLGAIIGDWV-DKNARKVQAQSLVVQN 100
Db 109 WSLVSVMF-----AVGGVSSFT-----GWIGRLGRVK-----AMLVVN 144
QY 101 VSVILCGIILMVFHLKHHELLTWYHGVLTSCYILITIANIANLASTATAITQDRIW 160
Db 145 VLSAGNLMGLAKMGSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
QY 161 VWAGEDRSKLANNNATIRRIDQLTNLAPWAVQIMTFGSPVIGCGFISG-----WNL--- 213
Db 187 YVSEVSPALRGALGTUHLQALVITGILI-----SQVLGLDFLGLNDELWPLLLG 235
QY 214 ---VSMCEVVELWKVQKTPALAVKAG-LKEBETELKQLNLKHDTPKPLEGTHLMGVK 269
Db 236 LSGVAALLQFFLLLLCPESPRYIKLGVVEAKSKLRLRGNC-----PMK----- 283
QY 270 DSNHLEHEQEPFCASQMAEPFTRDGMVSVYNQPVELA-----GMGLAFLYM 319
Db 284 --EIAEMEKEQEAASEKRVISIGLFSS-----SKYQAVIVALVMVQISQQSGINAFYVS 338
QY 320 T-----VLGPDCTITTYAYTQGLSGSILSILMGASA-----ITGIMG--- 356
Db 339 TNIFQAGVGPVVATIGVGWNTVF-----TVISVFLVEKAGRRSLFLAGLMGLI 390
QY 357 -TVAFWLRKCGLVRTGLISGLAQSLCLLCVISMVPM-----PG----- 395
Db 391 SAVAMT-----VGLV-----LLSQFAMWSYVSMVAIFLFFVFEVGGPIPMFVIAELFSQ 442
QY 396 -SPLDLSVSPFEDIRSRFIOG 415
Db 443 PRPAATAVAGFCNACNFIVG 463
RESULT 15
H2R2_RAT
ID_H2R2_RAT STANDARD; PRT; 358 AA.
AC P25102;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Histamine H2 receptor (H2R) (Gastric receptor 1).
GN HHR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:24:21 ; Search time 58 Seconds
(without alignments)
2540.483 Million cell updates/sec

Title: US-09-715-927-6

Perfect score: 2929

Sequence: 1 MTRAGDHNRRGCCGLADY.....ACGPDKEVRKENQANTSVV 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2929	100.0	571	4 Q9NP59	Q9np59 homo sapien
2	2921	99.7	571	4 Q8IVB2	Q8ivb2 homo sapien
3	2906	99.2	571	4 Q9NRL0	Q9nr10 homo sapien
4	2618.5	89.4	570	11 Q9JHI9	Q9jhi9 mus musculus
5	2612.5	89.2	570	11 Q9JKE4	Q9jke4 mus musculus
6	2600.5	88.8	570	11 Q923U9	Q923u9 rattus norv
7	1975.5	67.4	562	13 Q8JFW0	Q8jfw0 brachydanio
8	1971.5	67.3	562	13 Q919R3	Q919r3 brachydanio
9	1894.5	64.7	528	13 Q8AW23	Q8aw23 brachydanio
10	1513.5	51.7	436	13 Q8AW28	Q8aw28 brachydanio
11	1479.5	50.5	405	11 Q921C9	Q921c9 rattus norv
12	1360.5	46.4	305	11 Q8BME5	Q8bme5 mus musculus
13	859	29.3	167	4 Q9NUS1	Q9nus1 homo sapien
14	612.5	20.9	524	10 Q80905	Q80905 arabidopsis
15	564	19.3	498	10 Q9LZC8	Q9lzc8 arabidopsis
16	509.5	17.4	551	5 O44607	O44607 caenorhabdi

17	498	17.0	485	5 Q8IA95	Q8ia95 caenorhabdi
18	471.5	16.1	440	5 Q9BKT8	Q9bkt8 caenorhabdi
19	371	12.7	81	11 Q8BUM5	Q8bum5 mus musculu
20	301	10.3	598	10 Q8W4E7	Q8w4e7 arabidopsis
21	209.5	7.2	677	5 O01891	O01891 caenorhabdi
22	187.5	6.4	482	10 O04629	O04629 arabidopsis
23	164	5.6	37	11 Q9JIM9	Q9jim9 mus musculu
24	157	5.4	496	16 Q8G5T4	Q8g5t4 bifidobacte
25	136.5	4.7	713	16 Q983A3	Q983a3 rhizobium 1
26	135.5	4.6	412	16 Q8Y1P3	Q8y1p3 ralsstonia s
27	128	4.4	1287	5 Q9V6Z6	Q9v6z6 drosophila
28	123.5	4.2	409	16 Q97LU6	Q97lu6 clostridium
29	120.5	4.1	393	16 Q8G2C7	Q8g2c7 brucella su
30	120.5	4.1	763	16 Q8YFI5	Q8yfi5 brucella me
31	120	4.1	421	16 Q8G4V4	Q8g4v4 bifidobacte
32	120	4.1	481	5 Q9VNM1	Q9vnm1 drosophila
33	120	4.1	485	5 Q8S235	Q8sz35 drosophila
34	119.5	4.1	949	16 Q8FTZ6	Q8ftz6 corynebacte
35	119	4.1	537	16 Q8NLL2	Q8nll2 corynebacte
36	118	4.0	437	16 Q8YTP1	Q8ytp1 anabaena sp
37	117.5	4.0	783	10 Q65383	Q65383 arabidopsis
38	116	4.0	469	17 Q97VB9	Q97vb9 sulfolobus
39	115.5	3.9	418	16 Q9KG22	Q9kg22 bacillus ha
40	115.5	3.9	491	5 Q9V8N1	Q9v8n1 drosophila
41	115	3.9	408	16 Q97TP9	Q97tp9 clostridium
42	115	3.9	503	5 Q9XW19	Q9xw19 caenorhabdi
43	114.5	3.9	387	2 Q9KWU2	Q9kwu2 sphingomona
44	114.5	3.9	424	16 Q8NU27	Q8nu27 corynebacte
45	114	3.9	392	16 Q97QT3	Q97qt3 streptococc

ALIGNMENTS

RESULT 1

Q9NP59	PRELIMINARY;	PRT;	571 AA.
AC	Q9NP59;		
DT	01-OCT-2000 (TRENBLrel. 15, Created)		
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)		
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)		
DE	Iron-regulated transporter IREG1 (Ferroporlini) (Hypothetical protein)		
DE	(Solute carrier family 11 (Proton-coupled divalent metal ion transporters), member 3).		
DE	IREG1 OR FPNI OR DKFZP586J0624.		
GN	Homo sapiens (Human).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20337919; PubMed=10882071;		
RA	McKie A.T., Marciani P., Rolfs A., Brennan K., Wehr K., Barrow D.,		
RA	Miret S., Bomford A., Peters T.J., Farzaneh F., Hediger M.A.,		
RA	Hentze M.W., Simpson R.J.;		
RT	"A novel duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation.";		
RL	Mol. Cell 5:299-309(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RN	TISSUE=Placenta;		
RC	MEDLINE=20155474; PubMed=10693807;		
RX	Donovan A., Brownlie A., Zhou Y., Shepard J., Pratt S.J., Moynihan J.,		
RA	Paw B.H., Drejer A., Barut B., Zapata A., Law T.C., Brugnara C.,		
RA	Lux S.E. IV, Pinkus G.S., Pinkus J.L., Kingsley P.D., Palis J.,		
RA	Fleming M.D., Andrews N.C., Zon L.I.;		
RT	"Positional Cloning of Zebrafish Ferroportin1 Identifies a Conserved Vertebrate Iron Exporter.";		
RL	Nature 403:776-781(2000).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Uterus;		
RA	Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;		


```

RN SEQUENCE FROM N.A.
RP MEDLINE=20291023; PubMed=10828623;
RA Haile D.J.;
RT "Assignment of slc11a3 to mouse chromosome 1 band 1B and SLC11A3 to
RL human chromosome 2q32 by in situ hybridization.";
DR EMBL; AF215636; AAF80986.1; -
DR InterPro; IPR001993; Mitoch carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 571 AA; 62581 MW; 88707BFFC4A4B921 CRC64;

Query Match 99.2%; Score 2906; DB 4; Length 571;
Best Local Similarity 99.5%; Pred. No. 7.8e-215;
Matches 568; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MTRAGDHNRRGCCGSLADYLTSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Db 1 MTRAGDHNRRGCCGSLADYLTSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Qy 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
Db 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
Qy 121 LTMVHGWLTSYILITITIANIANLASTATAITIQRDWIVVWVAGEDRSKLANMATIRRI 180
Db 121 LTMVHGWLTSYILITITIANIANLASTATAITIQRDWIVVWVAGEDRSKLANMATIRRI 180
Qy 181 DOLTNILAPMAVGQIMTFGSPVIGCGFISGNWLSMCMVEYVLLWKVYQKTPALAVKAGLK 240
Db 181 DOLTNILAPMAVGQIMTFGSPVIGCGFISGNWLSMCMVEYVLLWKVYQKTPALAVKAGLK 240
Qy 241 BEETELKQLNLHKTPEKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 BEETELKQLNLHKTPEKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYYNQPVFLAGLAFYMTVLGDFCITTYAYTQGLSGSILSLMGASAITGIMGTAF 360
Db 301 SYYNQPVFLAGLAFYMTVLGDFCITTYAYTQGLSGSILSLMGASAITGIMGTAF 360
Qy 361 TWLRKCGLVRTGLISGLAQSLCILCVISVFMGSPDLDSVPSPEDIRSRFIOGESITP 420
Db 361 TWLRKCGLVRTGLISGLAQSLCILCVISVFMGSPDLDSVPSPEDIRSRFIOGESITP 420
Qy 421 TKIPETITIEIYMSNGSNANIVPETSPEVPTISVLSLLFAGVIAARIGLWSFDLTVTOLL 480
Db 421 TKIPETITIEIYMSNGSNANIVPETSPEVPTISVLSLLFAGVIAARIGLWSFDLTVTOLL 480
Qy 481 QENVIESERGIINGVQNSMYLLDLHLHFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMYLLDLHLHFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Qy 541 FAQNTLGNKLPACGPDPAKEVKENQANTSVV 571
Db 541 FAQNTLGNKLPACGPDPAKEVKENQANTSVV 571

RESULT 4
Q9JH19 PRELIMINARY; PRT; 570 AA.
ID Q9JH19
AC Q9JH19
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE SLC11A3 iron transporter (ferroportin) (Solute carrier family 39
DE (iron-regulated transporter) (ferroportin), member 1).
GN FPN1 OR SLC39A1.
OS Mus musculus. (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20347902; PubMed=10747949;
RA Abboud S., Haile D.J.;
RT "A Novel Mammalian Iron-regulated Protein Involved in Intracellular
RT Iron Metabolism";
RL J. Biol. Chem. 275:19906-19912 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20291023; PubMed=10828623;
RA Haile D.J.;
RT "Assignment of slc11a3 to mouse chromosome 1 band 1B and SLC11A3 to
RT human chromosome 2q32 by in situ hybridization.";
RL Cyogenet. Cell Genet. 88:328-329 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20155474; PubMed=10693807;
RA Donovan A., Brownlie A., Zhou Y., Shepard J., Pratt S.J., Moynihan J.,
RA Paw B.H., Drejer A., Barut B., Zapata A., Law T.C., Brugnara C.,
RA Lux S.E. IV, Pinkus G.S., Pinkus J.L., Kingsley P.D., Palis J.,
RA Fleming M.D., Andrews N.C., Zon L.I.;
RT "Positional Cloning of Zebrafish Ferroportin1 Identifies a Conserved
RT Vertebrate Iron Exporter";
RL Nature 403:776-781 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bokoff D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cdna collection.";
RL Nature 409:695-690 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF215637; AAF80987.1; -
DR EMBL; AF226613; AAF36696.1; -
DR EMBL; AK008700; BAB25840.1; -
DR EMBL; BC003438; AAH03438.1; -
DR MGD; MGI:1315204; SLC39A1.
DR InterPro; IPR001993; Mitoch carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 570 AA; 62702 MW; 7125CC6171394A0A CRC64;

Query Match 89.4%; Score 2618.5; DB 11; Length 570;
Best Local Similarity 90.2%; Pred. No. 9.4e-193;
Matches 518; Conservative 18; Mismatches 31; Indels 7; Gaps 3;

Qy 1 MTRAGDHNRRGCCGSLADYLTSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Db 1 MTRAGDHNRRGCCGSLADYLTSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Qy 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
Db 61 TAVYGLVAGSVLVGLGAIIGDWDKNAKLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
Qy 121 LTMVHGWLTSYILITITIANIANLASTATAITIQRDWIVVWVAGEDRSKLANMATIRRI 180

```

```
Db 121 LTMHGWLTVCYIIITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
Db 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
QY 241 EEBETELKQNLNKHDKTEPKLEGTHLMGVKDSNIHELEHEOEPTCASOMAEPPRTFRDGVW 300
Db 241 VBESELKQLTSPKDETPKLEGTHLMGEKDSNIHELEHEOEPTCASOMAEPPRTFRDGVW 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSILMGASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSILMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
Db 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
QY 421 -TKIPE--ITTEIYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLSFOLTVT 477
Db 421 TTIKIPETVFTTEHMSNMNSN---VHEMSTKPIPIVSLSLLFAGVIAARIGLSFOLTVT 476
QY 478 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVANGHLM 537
Db 477 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVANGHLM 536
QY 538 YFRPAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 537 YFRPAQNTLGNQIFVCGPDEKEVTDENQNTSVV 570

RESULT 5
Q9JKP4 PRELIMINARY; PRT; 570 AA.
ID Q9JKP4 AC Q9JKP4;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Iron-regulated transporter IREG1.
GN SLC39A1 OR IREG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20337919; PubMed=10882071;
RA McKie A.T., Marciani P., Rolfe A., Brennan K., Wehr K., Barrow D.,
RA Miret S., Bomford A., Peters T.J., Farzaneh F., Hediger M.A.,
RA Hentze M.W., Simpson R.J.;
RT "A novel duodenal iron-regulated transporter, IREG1, implicated in the
RT basolateral transfer of iron to the circulation."
RL Mol. Cell 5:299-309(2000).
DR EMBL; AF231120; AAF4329.1; -.
DR MGD; MGI:1315204; SLC39A1.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 570 AA; 62656 MW; C648A6C61C51EDCD CRC64;

Query Match 89.2%; Score 2612.5; DB 11; Length 570;
Best Local Similarity 90.1%; Pred. No. 2.7e-192;
Matches 517; Conservative 18; Mismatches 32; Indels 7; Gaps 3;

QY 1 MTRAGDHNRRGCCGLADYLTSAKFLLYLGHSLSTWGRMWHFAVSFVFLVLYGNSLL 60
Db 1 MTKARDQTHQEGCCGSLANLYTSAKFLLYLGHSLSTWGRMWHFAVSFVFLVLYGNSLL 60
QY 61 TAVYGLVWAGSVLVGLNIGDWDKQARLKVAQTSLVQVQNSVILCGIILMVFLKHKL 120
Db 61 TAVYGLVWAGSVLVGLNIGDWDKQARLKVAQTSLVQVQNSVILCGIILMVFLKHKL 120
QY 121 TAVYGLVWAGSVLVGLNIGDWDKQARLKVAQTSLVQVQNSVILCGIILMVFLKHKL 120
Db 61 TAVYGLVWAGSVLVGLNIGDWDKQARLKVAQTSLVQVQNSVILCGIILMVFLKHKL 120
QY 121 LTMHGWLTVCYIIITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
```

```
Db 121 LTMHGWLTVCYIIITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
Db 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
QY 241 EEBETELKQNLNKHDKTEPKLEGTHLMGVKDSNIHELEHEOEPTCASOMAEPPRTFRDGVW 300
Db 241 VBESELKQLTSPKDETPKLEGTHLMGEKDSNIHELEHEOEPTCASOMAEPPRTFRDGVW 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSILMGASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSILMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
Db 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
QY 421 -TKIPE--ITTEIYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLSFOLTVT 477
Db 421 TTIKIPETVFTTEHMSNMNSN---VHEMSTKPIPIVSLSLLFAGVIAARIGLSFOLTVT 476
QY 478 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVANGHLM 537
Db 477 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVANGHLM 536
QY 538 YFRPAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 537 YFRPAQNTLGNQIFVCGPDEKEVTDENQNTSVV 570

RESULT 6
Q923U9 PRELIMINARY; PRT; 570 AA.
ID Q923U9 AC Q923U9;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Ferroportin 1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Yeh K.-Y., Yeh M., Glass J.;
RT "Iron induces ferroportin 1 (Fp1) clustering and redistribution in
RT rat intestinal epithelial cells."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394785; AAK77858.2; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 570 AA; 62586 MW; 766786F22F054E94 CRC64;

Query Match 88.8%; Score 2600.5; DB 11; Length 570;
Best Local Similarity 89.9%; Pred. No. 2.3e-191;
Matches 514; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

QY 1 MTRAGDHNRRGCCGLADYLTSAKFLLYLGHSLSTWGRMWHFAVSFVFLVLYGNSLL 60
Db 1 MTKSRDQTHQEGCCGSLANLYTSAKFLLYLGHSLSTWGRMWHFAVSFVFLVLYGNSLL 60
QY 61 TAVYGLVWAGSVLVGLNIGDWDKQARLKVAQTSLVQVQNSVILCGIILMVFLKHKL 120
Db 61 TAVYGLVWAGSVLVGLNIGDWDKQARLKVAQTSLVQVQNSVILCGIILMVFLKHKL 120
QY 121 LTMHGWLTVCYIIITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
Db 121 LTMHGWLTVCYIIITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
```

```
Db 181 DLTNLTAPMAVQIMTGPVIGCGFTSGNVLVSMCVYELLWKVYQKTPALAVKALK 240
Qy 241 EETELKQNLNKHDKTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTPDCGW 300
Db 241 VEESLQKLTSPKDEPKPLEGTHLMGVKDSNIHELEHQEPTCASQIAEPFRTPDCGW 300
Qy 301 SYNOPVFLAGMLAFLYMTVLGFCITTCGYAYTQGLSGSILSLMGLASAITGIMGTVA 360
Db 301 SYNOPVFLAGMLAFLYMTVLGFCITTCGYAYTQGLSGSILSLMGLASAITGIMGTVA 360
Qy 361 TWLRKCGVRLTGLSGLAQLSCLILCVISVFMPSGLDLSVPSPEDIRSRFIQGESI-T 419
Db 361 TWLRKCGVRLTGLSGLAQLSCLILCVISVFMPSGLDLSVPSPEDIRSRFIHEAVSS 420
Qy 420 PTKIPEITTEIYMSNGSNANIIVPETSPIISVSLFAGVIAARIGLWSFDLTVTQ 479
Db 421 TTKIPE--TEMLMSNVNVTVMHSTKSPISVSLFAGVIAARIGLWSFDLTVTQ 478
Qy 480 LOENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLSVSVFVAMGHIMYF 539
Db 479 LOENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLSVSVFVAMGHIMYF 538
Qy 540 RFAQNTLGNKLPACGPDAKEVRKENQANTSVV 571
Db 539 RFAQNTLGNKLPACGPDAKEVTDQSPTNSV 570
```

```
RESULT 7
Q8JFWO PRELIMINARY; PRT; 562 AA.
ID Q8JFWO
AC Q8JFWO;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE SI:dz211013.7 (Solute carrier family 39 (Iron-regulated transporter), member 1).
DE member 1).
GN SLC39A1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SLC39A1.
RA Babbage A.;
RP SEQUENCE FROM N.A.
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL591593; CAD3474.1; -.
SQ SEQUENCE 562 AA; 61761 MW; 35CB30971553718F CRC64;
```

```
Query Match 67.4%; Score 1975.5; DB 13; Length 562;
Best Local Similarity 71.0%; Pred. No. 2.3e-143;
Matches 394; Conservative 64; Mismatches 72; Indels 25; Gaps 10;

Qy 14 CGSLADYLTSAKFLYLGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 73
Db 11 CERPREFFKSAKFLIYVGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 70
Qy 74 VLGAIGDWVDKNAKLKVAQTSLVQVNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 133
Db 71 LLGAIGDWVDKNAKLKVAQTSLVQVNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 130
Qy 134 ILIITIANIANLASTATAITQRDWIVVAGDSRKLANNATRIIDLTNIIAPMAVG 193
Db 131 IMVISTIANIANLASTAMISITIQRDWIVVAGDSRKLANNATRIIDLTNIIAPMAVG 190
Qy 194 QIMTFGSPVIGCGFTSGNVLVSMCVYELLWKVYQKTPALAVKAGLKE-EETELKQNLN 252
Db 191 QIMAFGSHFICGFTSGNVLVSMCVYELLWKVYQKTPALAVKAGLKE-EETELKQNLN 250
Qy 253 KD---TEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTPDCGWVSYNQPVFL 309
Db 251 KEIGNTE-SPVEASQLM-----TESGPKKDTGCCQMAEPFRTPDCGWVSYNQPVFL 303
```

```
Qy 310 AGMLAFLYMTVLGFCITTCGYAYTQGLSGSILSLMGLASAITGIMGTVAFTWLRKCGL 369
Db 304 AGMLAFLYMTVLGFCITTCGYAYTQGLSGSILSLMGLASAVSGICGTVAFTWLRKCGL 363
Qy 370 VRTGLISGLAQLSCLILCVISVFMPSGLDLSVPSPEDIRSRFIQGESI-----IPTTKIPE 425
Db 364 IRTGFIAGVTQLSCLILCVASVAPGSPFDLSVPSPFEEV-LRHLFGDSGSLRESPTFTP- 421
Qy 426 ITTEIYMSNGSNANI-VPETSPIISVSLFAGVIAARIGLWSFDLTVTQLOQNV 484
Db 422 -TTTPPI-----QANVTVFEEAPPVESYMSVGLLPAGVIAARVGLWSFDLTVTQIQNV 475
Qy 485 IESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLSVSVFVAMGHIMYFRAQN 544
Db 476 IESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLSVSVFVAMGHIMYFRAQN 535
Qy 545 TLGNKLPFA-CGPDAK 558
Db 536 SLGSRLEFLFCSPQK 550
```

RESULT 8

```
ID Q919R3 PRELIMINARY; PRT; 562 AA.
AC Q919R3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Ferroportin1.
GN SLC39A1 OR FPN1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SLC39A1.
RA SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RC MEDLINE=20155474; PubMed=10693807;
RA Donovan A., Brownlie A., Zhou Y., Shepard J., Pratt S.J., Moynihan J.,
RA Paw B.H., Drejer A., Barut B., Zapata A., Law T.C., Brugnara C.,
RA Lux S.E. IV, Pinkus G.S., Pinkus J.L., Kingsley P.D., Palis J.,
RA Fleming M.D., Andrews N.C., Zon L.I.;
RT "Positional Cloning of Zebrafish Ferroportin1 Identifies a Conserved
RT Vertebrate Iron Exporter.";
RL Nature 403:776-781(2000).
DR EMBL; AF226612; AAF36695.1; -.
DR ZFIN; ZDB-GENE-000511-8; slc39a1.
SQ SEQUENCE 562 AA; 61760 MW; 38A6838E2107FBFAF CRC64;
```

```
Query Match 67.3%; Score 1971.5; DB 13; Length 562;
Best Local Similarity 70.8%; Pred. No. 4.7e-143;
Matches 393; Conservative 65; Mismatches 72; Indels 25; Gaps 10;

Qy 14 CGSLADYLTSAKFLYLGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 73
Db 11 CERPREFFKSAKFLIYVGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 70
Qy 74 VLGAIGDWVDKNAKLKVAQTSLVQVNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 133
Db 71 LLGAIGDWVDKNAKLKVAQTSLVQVNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 130
Qy 134 ILIITIANIANLASTATAITQRDWIVVAGDSRKLANNATRIIDLTNIIAPMAVG 193
Db 131 IMVISTIANIANLASTAMISITIQRDWIVVAGDSRKLANNATRIIDLTNIIAPMAVG 190
Qy 194 QIMTFGSPVIGCGFTSGNVLVSMCVYELLWKVYQKTPALAVKAGLKE-EETELKQNLN 252
Db 191 QIMAFGSHFICGFTSGNVLVSMCVYELLWKVYQKTPALAVKAGLKE-EETELKQNLN 250
Qy 253 KD---TEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTPDCGWVSYNQPVFL 309
Db 251 KEIGNTE-SPVEASQLM-----TESGPKKDTGCCQMAEPFRTPDCGWVSYNQPVFL 303
```

```
Db 251 KEIGNTE-SPVEASQLM-----TESSEPKKDTGCCYQMAEPIRTFKDGMVAYVYNQSIFF 303
QY 310 AGMGLAFYMTVLGDFDCTTGYAYTQGLSGSILSMGASAITGIMGTVAFTWLRKCGL 369
Db 304 AGMSLAFYMTVLGDFDCTTGYAYTQGLSGSVLSLLMGASAVSGICGTVAFTWIRKCGL 363
QY 370 VRTGLISGLAQLSCLILCVISVFMPSGLDLSVPPEDIRSRFTQGES-----ITPTKIPE 425
Db 364 IRTGFTAGVTQLSCLTLCVASVFAPGSPFDLSVPFKEV-LRHLFGDSGLSRSPFTIP- 421
QY 426 ITTEIYMSNGSNSANI-VPETSPESVPIISVSLFAGVIAARIGLMSFDTLTQTLLQENV 484
Db 422 -TTEPPI-----QANVTVFEEAPPVESYMSVGLLFAGVIAARVGLMSFDTLTQTLLQENV 475
QY 485 IESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFRPAQN 544
Db 476 IESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFRPAQN 535
QY 545 TLGNKLFPA-CGPDAK 558
Db 536 SLGSLRFLFCSPEQK 550

RESULT 9
Q8AW23 PRELIMINARY; PRT; 528 AA.
ID Q8AW23
AC Q8AW23;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SI:d218H3.5 (Solute carrier family 39 (Iron-regulated transporter),
member 1) (Fragment).
GN SLC39A1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimberley A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL672145; CAD58776.1; -
FT NON TER 1
SQ SEQUENCE 528 AA; 57748 MW; 07ECE661DBE18851 CRC64;

Query Match 64.7%; Score 1894.5; DB 13; Length 528;
Best Local Similarity 71.8%; Pred. No. 3.5e-137;
Matches 381; Conservative 59; Mismatches 66; Indels 25; Gaps 10;

QY 38 GDRMHPAVSVFLVELYGNLSLLTAYGLVAGSVLVGLGAIIGDWVDKRNARKVAOTSLV 97
Db 1 GDRMHPAVSVFLVELYGNLSLLTAYGLVAGSVLVGLGAIIGDWVDKRNARKVAOTSLV 60
QY 98 VQNSVILCGIILMVLFLHKEHLLTMVHGWLTVSCVLIITIANIANLASTATAITQD 157
Db 61 VQNSVILCGIILMVLFLHKEHLLTMVHGWLTVSCVLIITIANIANLASTATAITQD 120
QY 158 WIVVAGEDRSKLANMATIRRIDQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLSMC 217
Db 121 WIVVAGEDRSKLANMATIRRIDQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLSMC 180
QY 218 VEYVLWKVYQKTPALAVKAGLKE-BETELKQNLHKO-----TEPKPLEGTHLMGVKDSNI 273
Db 181 LEYFLWKVYQKTPALAVKAGLKE-BETELKQNLHKO-----TEPKPLEGTHLMGVKDSNI 233
QY 274 HELEHQEPTCASOMAEPRFTFRDGMVSYNQVPELAGMGLAFYMTVLGDFDCTTGYAY 333
Db 234 ESSETKKDTGCCYQMAEPIRTFKDGMVAYVYNQSIFFAGMSLAFYMTVLGDFDCTTGYAY 293
QY 334 TQGLSGSILSMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLILCVISVFM 393
Db 294 TQGLSGSVLSLLMGASAVSGICGTVAFTWIRKCGLIRTGFTAGVTQLSCLTLCVASVFA 353

Query Match 51.7%; Score 1513.5; DB 13; Length 436;
Best Local Similarity 69.7%; Pred. No. 5.1e-108;
Matches 306; Conservative 51; Mismatches 57; Indels 25; Gaps 10;

QY 130 TSCYIIITIANIANLASTATAITQDWMVIVVAGEDRSKLANMATIRRIDQLTNILAP 189
Db 1 TTCYIMVSIANIANLASTAMSIITQDWMVIVVAGEDRSKLANMATIRRIDQLTNILAP 60
QY 190 MAVGQIMTFGSPVIGCGFISGWNLSMCVVEYVLLMKVYQKTPALAVKAGLKE-BETELKQ 248
Db 61 MLVGQIMAFGSHFICGCGFISGWNLSMCLEYLEYLLMKVYQKTPALAVKAGLKE-BETELKQ 120
QY 249 LNLHKO-----TEPKPLEGTHLMGVKDSNIHELEHQEPTCASOMAEPRFTFRDGMVSYNQ 305
Db 121 LNLHKO-----TEPKPLEGTHLMGVKDSNIHELEHQEPTCASOMAEPRFTFRDGMVSYNQ 173
QY 306 PVFLAGMGLAFYMTVLGDFDCTTGYAYTQGLSGSILSMGASAITGIMGTVAFTWLR 365
Db 174 SIFFAGMSLAFYMTVLGDFDCTTGYAYTQGLSGSVLSLLMGASAVSGICGTVAFTWIRK 233
QY 366 KCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPEDIRSRFTQGES-----ITPT 421
Db 234 KCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPEDIRSRFTQGES-----ITPT 292
QY 422 KPEITTEIYMSNGSNSANI-VPETSPESVPIISVSLFAGVIAARIGLMSFDTLTQTLL 480
Db 293 FIP--TTEPPI-----QANVTVFEEAPPVESYMSVGLLFAGVIAARVGLMSFDTLTQTLL 345
QY 481 QENVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFR 540
Db 346 QENVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFR 405
QY 541 FAQNTLGNKLFPA-CGPDAK 558
Db 406 FAYKSLGSLRFLFCSPEQK 424
```



```
RESULT 11
Q921C9 PRELIMINARY; PRT; 405 AA.
AC Q921C9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Cell adhesion regulator.
GN CARL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang X.Z.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U76714; AAD00260.1; -.
SQ SEQUENCE 405 AA; 45145 MW; ABD5916C43AA4DA CRC64;

Query Match 50.5%; Score 1479.5; DB 11; Length 405;
Best Local Similarity 78.4%; Pred. No. 1.9e-105;
Matches 298; Conservative 15; Mismatches 36; Indels 31; Gaps 4;

Qy 1 MTRAGDHNRQRCGCCSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSVFLVELYXNSLL 60
Db 1 MTKSRDQTHQSGCCSLANYLTSAKFLLYLGHSLSTWGDMMHFAVSVFLVELYXNSLL 60

Qy 61 TAVYGLVAGSVLVGALIGDWDKNARKVAQTSLVVQNVSVILCGIILMMVFLKHKL 120
Db 61 TAVYGLVAGSVLVGALIGDWDKNARKVAQTSLVVQNVSVILCGIILMMVFLKHKL 120

Qy 121 LTMVHGWLTCYLIITIANIANLSTATAITTDWIVVWVAGDSKLANNNATIRI 180
Db 121 LNMTHGWLTCYLIITIANIANLSTATAITTDWIVVWVAGDSKLANNNATIRI 180

Qy 181 DQNLNAPMAGVQIMTFSPVIGCGFTSGWNLVSMCVYVLLMKVQKTPALAVKAGLK 240
Db 181 DQNLNAPMAGVQIMTFSPVIGCGFTSGWNLVSMCVYVLLMKVQKTPALAVKAGLK 240

Qy 241 BEETELKQLNKHKTEPKPLEGTHLMGVKDSNIHELEHQBPPTCASQWABPFRTRDGV 300
Db 241 VEESLQKLTSPKTEPKPLEGTHLMGEKDSNIHELEHQBPPTCASQIABPFRTRDGV 300

Qy 301 SYVNPQVFLAGWGLAF-----LYMTVLGFDCTTG--YATQGLSGSILMGAS 349
Db 301 SYVNPQVFLAGWGLAF-----LYMTVLGFDCTTG--YATQGLSGSILMGAS 349

Qy 350 AITGIMGTVAFTLRRKCGL 369
Db 352 -----NWNNGNCGL 360

RESULT 12
Q8BME5 PRELIMINARY; PRT; 305 AA.
AC Q8BME5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Solute carrier family 39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Mesonephros;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK032732; BAC28001.1; -.
SQ SEQUENCE 305 AA; 33574 MW; EC9FE6264AF4D887 CRC64;

Query Match 46.4%; Score 1360.5; DB 11; Length 305;
Best Local Similarity 88.0%; Pred. No. 1.8e-96;
Matches 272; Conservative 11; Mismatches 19; Indels 7; Gaps 3;

Qy 266 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSYVNPQVFLAGWGLAFLYMTVLGFD 325
Db 1 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSYVNPQVFLAGWGLAFLYMTVLGFD 60

Qy 326 CITTGAYTQGLSGSILSLILMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLI 385
Db 61 CITTGAYTQGLSGSILSLILMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLI 120

Qy 386 LCVISVPMGSPDLDSVSPEDIRSRFIOGESITP-TKIFE--ITTEIYMSNGSNSANIV 442
Db 121 LCVISVPMGSPDLDSVSPEDIRSRFVNVEPVSPTTKIPETVFTTEHMSNMSN---V 176

Qy 443 PETSPEVPIISVSLIFAGVIAAIGLWSFDLTVTQLQENVIESERGIINGVONSNNYL 502
Db 177 HEMSTKPIPIVSVSLIFAGVIAAIGLWSFDLTVTQLQENVIESERGIINGVONSNNYL 236

Qy 503 LDLLHFTMVLANPEAFGLLVLISVSFVANGHIMYFRFAQNTLGNKLFACGPDKEVRK 562
Db 237 LDLLHFTMVLANPEAFGLLVLISVSFVANGHLMYFRFAQNTLGNQIFVCGPDEKEVTD 296

Qy 563 ENQNTSVV 571
Db 297 ENQNTSVV 305

RESULT 13
Q9NUS1 PRELIMINARY; PRT; 167 AA.
AC Q9NUS1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein FLJ11176.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
RA Wagaatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK002038; BAA92049.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein.
SQ SEQUENCE 167 AA; 18252 MW; 8D391F5356733660 CRC64;

Query Match 29.3%; Score 859; DB 4; Length 167;
Best Local Similarity 98.8%; Pred. No. 2.8e-58;
Matches 165; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 266 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSYVNPQVFLAGWGLAFLYMTVLGFD 325
Db 1 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSYVNPQVFLAGWGLAFLYMTVLGFD 60

Qy 326 CITTGAYTQGLSGSILSLILMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLI 385
```

```
Db 61 CTTGAYTQGLSGSILSLMGASATGNGTAVFTWLRKCGLVRTGLISGLAQLSCLL 120
QY 386 LCVISVMPGSPDLVSPEDIRSRFIQGESITPTKIPITTEIYM 432
Db 121 LCVISVMPGSPDLVSPEDIRSRFIQGESITPTKIPITTEIHW 167

RESULT 14
ID 080905 PRELIMINARY; PRT; 524 AA.
AC 080905;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE At2g38460 protein.
GN At2G38460.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Frazer C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004683; AAC28758.1; -.
SQ SEQUENCE 524 AA; 58204 MW; 781C6C0A697FADB1 CRC64;

Query Match 20.9%; Score 612.5; DB 10; Length 524;
Best Local Similarity 29.1%; Pred. No. 1.1e-38;
Matches 155; Conservative 95; Mismatches 179; Indels 103; Gaps 9;

QY 28 LYLGHSLSTWGDMMHFAVSFVFLVLYGNSLLLTAVYGLVVGSLVGLGAILGDWYDKNA 87
Db 41 LYVGYFLRWSARTWFSVALYMIHLWPNLSLLAAIYGAIESGTALFGPIVGQWEGMD 100
QY 88 RLKVAQTSLVQNVSVILCG---IILMMVFLHKHELLTMYHGWLTSCYILITITIANIAN 144
Db 101 YVKVLRLLWLFQNLSTYIAGGAVIKLLVSDLSKRNLPV-----FALLVLTNLAG 151
QY 145 -----LASTATAITIQDWIVVVA-GEDRSKLANNTATIRIDQLTNILAPMAVGQIMTFG 199
Db 152 AIGVLTSLAGTILIERDWAVMSEGHPPAVLTQNSVIRGIDLSKLLSPVITGLIISFV 211
QY 200 SPVIGCGFISGWNLVSMCVYVLLWKVYQKTPALAVKAGLKEETELKQNLHKDTEPKP 259
Db 212 SLKASAITFAAWATTATWVYEWLFISVGSVPAT-----RSNERRILSRTKQVEGRDAP 267
QY 260 LEGTHLMGVKDSNIHELEHQEPTCASQMAEPRT-----FRDGMVSYNQP 306
Db 268 VSVSIVPGTEEG-----YTGNP-----PSRTGILVDRMSKSSFVGAWRIVFNQE 313
QY 307 VFLAGGLAFLYMTVLGFCDCITTYAYTQGLSSILSLMGASATGNGTAVFTWLRK 366
Db 314 VLPFGVSLALLFTVLSFGTMTATLQWEGIPTYYIIGIRGISATVGLAATVLYPLMQSR 373
QY 367 CGLVRTGLISGLAQLSCLILCVISVMPGSPDLVSPEDIRSRFIQGESITPTKIP 426
```

```
Db 374 LSTLTGLWSFWSQSCLLVCVGSIVWKDK----- 404
QY 427 TTEIYMSNGSNSANIVPETSPEPVIIISVSLFAGVIAARIGLWSFDLTVTQLLENVIE 486
Db 405 -----IASYMLMAGVAASRLGLMFMFLAVIQMQDLVSE 438
QY 487 SERGIINGVQNSMNYLLDLLLHFIMVILAPNPEAFGLLVLSVSVFAMGHIMY 538
Db 439 SDRCVGVGVQNSLQSLDLMAYLLGLIIVSNPKDFWILLISFSTVSLAGMLY 490

RESULT 15
QYLCZ8 PRELIMINARY; PRT; 498 AA.
ID Q9LZC8
AC Q9LZC8;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE Transporter like protein.
GN F12B4370..
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162751; CAB83320.1; -.
SQ SEQUENCE 498 AA; 55504 MW; D8A4C5F61827998D CRC64;

Query Match 19.3%; Score 564; DB 10; Length 498;
Best Local Similarity 27.5%; Pred. No. 5.3e-35;
Matches 153; Conservative 102; Mismatches 175; Indels 126; Gaps 12;

QY 28 LYLGHSLSTWGDMMHFAVSFVFLVLYGNSLLLTAVYGLVVGSLVGLGAILGDWYDKNA 87
Db 37 LYLGYFLRWSARTWFSVALYMIHLWPNLSLELTAMYGVSEGSATLFGPIVGQWIDGMS 96
QY 88 RLKVAQTSLVQNVSVILCG---IILMMVFLHKHELLTMYHGWLTSCYILITITIANIAN 144
Db 97 YVKVLRLLWLTQNLSTYIAGGAVVALLVVPDLKSNQFPV-----FATLVLTNLSG 147
QY 145 -----LASTATAITIQDWIVVVA-GEDRSKLANNTATIRIDQLTNILAPMAVGQIMTFG 199
Db 148 AIGVLTSLAGTILIERDWAVMSEGHSPAVLTQNSVIRGIDLSKLLSPVITGLIISFV 207
QY 200 SPVIGCGFISGWNLVSMCVYVLLWKVYQKTPALA---VKAGLKEEETELKQNL----- 250
Db 208 SLRASAITFAAWATTATWVYEWLFISVNGVPAIVQSDERRSLRSQSQAETDSASSFY 267
QY 251 ---LHKDTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPRTFRDGMVSYNQP 307
Db 268 VPLLHEESYRNTQS-----RSRILRILRISE-----SSFVSARNVNLQEI 310
QY 308 FLAGGLAFLYMTVLGFCDCITTYAYTQGLSSILSLMGASATGNGTAVFTWLRK 367
Db 311 VLPFGVSLALLFTVLSFGTMTATLQWEGIPTYYIIGIRGISAGVGLAATVLYPLMQSR 370
QY 368 GLVVRTGLISGLAQLSCLILCVISVMPGSPDLVSPEDIRSRFIQGESITPTKIP 427
Db 371 SPRTGVWSFSQK-----EKIAS----- 389
QY 428 TTEIYMSNGSNSANIVPETSPEPVIIISVSLFAGVIAARIGLWSFDLTVTQLLENVIE 487
Db 390 ---YM-----LMAGVAASRLGLMFMFLAVIQMQDLVSE 421
```

```

Qy 488 ERGLINGVONSMTYLLDLLHFMVILAPNPEAFGLLVISVSFVAMGHIMY----- 538
Db 422 DRCVGVGVSQSLDLMANLLGIIVSNPKDFWMLTLISFATVSLAGILYTHLYIRK 481
Qy 539 --FPAQNTLGNKLFA 552
Db 482 HLFHLEKIPLNNFFA 497

```

Search completed: November 5, 2003, 18:30:50
Job time : 62 secs

THIS PAGE BLANK (USPTO)